KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY

INSITUTE OF DISTANCE LEARNING

COLLEGE OF SCIENCE

DEPARTMENT OF BIOLOGICAL SCIENCE

KNUST

PREDICTING BOD LEVELS OF WASTEWATER WITH NEURAL NETWORK TIME SERIES

By

IBRAHIM ABDALLAH, BSC (PHYSICS)

JUNE, 2014

PREDICTING BOD LEVELS OF WASTEWATER WITH NEURAL NETWORK TIME SERIES

KNUST

By

ABDALLAH, IBRAHIM

(PG4166310)

A THESIS SUBMITTED TO THE INSTITUTE OF DISTANCE LEARNING, KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN ENVIRONMENTAL SCIENCE

JUNE, 2014

DECLARATION

I hereby declare that this submission is my own work towards the Masters of Science in Environmental Science and that, to the best of my knowledge, it contains no material previously published by another person nor material which has been accepted for the award of any other degree of the University, except where due acknowledgement has been made in the text.

KNUST

Ibrahim Abdallah, PG4166310	………………..	…………………….
Student's Name & ID	Signature	Date

Certified by:

E. Owusu Ansah	………………	…………………
Supervisor's Name	Signature	Date

Certified by:

Rev. S. Akyeampong	…………………	…………………
Head of Department	Signature	Date

NYANSAPO W J SANE NO BADWENIMA

DEDICATION

In memory of my late parents;

My Father, Mr. Abukari Wavie Yomoh and My Mother Madam Mariam Hapommie

KNUST

ACKNOWLEDGEMENTS

In writing this thesis, I was helped by various people-especially by E. Owusu Ansah, my supervisor, who first suggested that I should write on this topic. He also supported me with research materials and most importantly his time and advice. Thanks to all my fellow students and friends who helped in explaining various statistical techniques and modeling concepts to me, and I am particularly grateful for their willingness to repeat the same thing multiple times if necessary. My friend, Godwin, his wife, Gifty and daughter, Glyniss who I shared their room with were great people to spend time with and I am especially grateful to them for their friendship, help and support throughout the year. I appreciate your effort and my God bless you.

ABSTRACT

The quality of treated wastewater has always been an important issue, but it becomes even more critical as human populations increase. An accurate well-timed measurement of quality variables is essential to the successful monitoring and controlling of wastewater treatment systems. Unfortunately, current ability to monitor and control effluent quality from a wastewater treatment process is primitive. Control is difficult because wastewater treatment consists of complex multivariate processes with nonlinear relationships and time varying dynamics. Because the measurements of these variables are difficult and often involve large time delays, there is a critical need for forecasting models that are effective in predicting wastewater effluent quality. In this paper, predictive models based on artificial neural networks are presented. Water quality measurements and process data from an urban wastewater treatment plant were used to develop models to predict biochemical oxygen demand (BOD). The results provide evidence that nonlinear neural network time series models achieve accurate forecast of wastewater effluent quality.

KNUST
NYANSAPƆ WƆSANE NO BADWENMA

TABLE OF CONTENTS

Content Page

KNUST
NYANSAPO W) SANE NO BADWENIMA

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADI	Acceptable Daily Intake
AMN	Ammoniacal Nitrogen
ANN	Artificial Neural Network
BOD	Biochemical Oxygen Demand
COD	Chemical Oxygen Demand
Cu	Copper
EC	Electrical Conductivity
FAO	Food and Agricultural Organization
FIR	Finite Impulse Response
GMNN	Gamma Memory Neural Network
KJN	Kjeldahl's Nitrogen
LR	Learning Rate
MAPE	Mean Average Percentage Error
Mg	Magnesium
MLP	Multi-Layer Perceptron
Mn	Manganese
MSE	Mean Square Error
N	Nitrogen
NH	Number of Hidden Layers
P	Phosphorus
SAR	Sodium Adsorption Ratio
SOFM	Self-Organising Feature Maps

KNUST

TDNN	Time Delay Neural Network
TDS	Total Dissolved Solids
TOC	Total Organic Carbon
TSS	Total Suspended Solids
UNO	United Nations Organization
US.EPA	United States Environmental Protection Agency
WHO	World Health Organization
Zn	Zinc

KNUST

CHAPTER ONE

INTRODUCTION

1.1 Background to the Study

It is a common practice today in many nations to discharge treated wastewater into the ecosystem. Because wastewater treatment is a costly process, wastewater is not typically purified to be safe for human consumption. The aquatic environment is often utilized to help in the "cleansing" process. As populations increase, higher quantities of treated wastewater are discharged into the environment which can overload ecosystems, leading to unsafe conditions for humans and animals. This has actually happened in the Narragansett Bay in Rhode Island, USA, which has several wastewater treatment facilities discharging along its shores. The bay has been closed to commercial fishing and recreation on many occasions when residents have become ill from eating food caught in the bay or from exposure to the water. Unfortunately, it is currently very difficult to predict when a body of water will become unsafe. Thus, public health is under a constant threat from potentially unclean waters (Yeung and Yung, 1999). The present ability to monitor and control the effluent quality from a wastewater treatment process is described by researchers as primitive and notoriously difficult (Wen and Vassiliadis, 1998; Boger, 1997). Wastewater treatment has been characterized as a complex multivariate process with highly variable inputs, nonlinear time varying dynamics, and an auto correlated time series structure that is subject to large disturbances (Lindberg, 1997).

It is of great importance in water quality control that the amount of organic matter present in the system be known and that the quantity of oxygen required for its stabilisation be determined. Over the years, different physico-chemical tests have been developed to determine the organic and inorganic content of wastewater (Metcalf and Eddy, 1995). In general, these tests may be divided into those used to measure gross concentrations of organic matter greater than 1mg/l and those used to measure trace concentrations in the range of 10^{-6} to 10^{-3} mg/l. Laboratory methods commonly used today to measure the gross amount of organic matter (greater than 1 mg/l in wastewater includes the following: Biochemical Oxygen Demand (BOD_5), Chemical Oxygen Demand (COD) and Total Organic Carbon (TOC).

These three parameters are used in wastewater treatment operations to estimate the influent and effluent characteristics and treatment efficiency. The use of TOC as an analytical parameter has become more common in recent years especially for the treatment of industrial wastewater. Partly, this is due to the fact that the TOC determinations can be carried out in triplicate within minutes compared with the five days required for the BOD_5 test (Sawyer *et al.*, 1994). Apart from these, the easily measurable parameters for any industrial wastewater include indices like Total Suspended Solids (TSS), Total Dissolved Solids (TDS), Phenol concentration, Ammoniacal Nitrogen (AMN) and Kjeldahl's Nitrogen (KJN) (Metcalf and Eddy, 1995). A review of the existing literature in this field reveals that correlation among these parameters seldom exists. It could be difficult to understand the dynamics of the relationship between these parameters because they primarily depend on the process of the target industry, raw material/by-product composition, composition of chemicals discharged in wastewater and thus their nonlinear relationship makes universal generalization difficult.

In addition, wastewater measurement systems can be unreliable and require lead times as long as five days to measure biochemical oxygen demand (BOD). The effluent BOD level is an important indication of the quantity of oxygen that will be depleted from the aquatic environment. Recently it has become possible to exercise more effective control of wastewater processes based on (1) the development of advanced forecasting models, (2) more accurate and timely measurement of process variables (Parkinson, 1998), and (3) wireless technology that provides a real time view of the treatment process (Cheek and Wilkes,1994).

1.2 Statement of the Problem

To measure the effectiveness of biological treatment, the test of the biochemical oxygen demand (BOD_5 or simply BOD) is widely used in the area of wastewater treatment. This standard method for the examination of water and wastewater uses a five day sample-incubation period at $20^{o}C$ to get the value of the BOD from laboratory tests. Therefore the BOD cannot be determined until 5 days after the sample has been taken. This 5 day delay makes it impossible for the operators to take instant action to control BOD fluctuations. Therefore, developing an effective BOD model for the timely prediction of treatment results has become a very pressing task in the optimization of the wastewater treatment. Neural networks have been applied to solve and predict problems related to the following; biodegradation kinetics of organic compounds (Shuurmann and Muller, 1994), estimating optimum alum doses in water treatment (Maier *et al*., 2004) and long term tidal waves (Lee, 2004).

1.3 Objectives of the study

The objectives set for this research were;

1. to determine the daily BOD levels of a wastewater plant
2. to use the determined BOD levels together with Time Delay Neural Networks (TDNN) to come out with a model that can predict future BOD levels in wastewater.
3. to test the predictive power of the model obtained by comparing forecasted values to actual BOD values.

1.4 Research questions

Questions addressed by this research to achieve its set objectives were;

1. Do wastewater have a significant amount of BOD levels?
2. What model can be developed to predict BOD levels in waste water?
3. Is the predictive power of this model strong enough?

1.5 Significance of the Study

There are many characteristics to monitor an on-site wastewater treatment system's performance. They vary from something as simple as checking for sewage on the surface, to complicated laboratory analysis. The cost and the required amount of sample vary from laboratory to laboratory. These laboratories need to be contacted in advance prior to dropping off samples. In choosing a laboratory to perform analysis of wastewater characteristics requires certified laboratories that use standard procedures. A list of certified laboratories is maintained by most developed countries. Minnesota Valley

Testing Labs. Inc is one of the accredited laboratories in the US (source: http://www.health.state.mn.us/divs/phl/cert/allcertlabs.html). The situation is however different in our part of the world where these facilities are hardly available. The cost, time and resources for laboratory analysis are too high. This result in poor or no monitoring of the levels of these wastewater characteristics which eventually end up in the environment and the negative implication are enormous. A model developed to predict the levels of these wastewater characteristics is therefore not far from the solutions to these problems. It is therefore justified that this project tries to use TDNN to predict the wastewater BOD levels of wastewater treatment systems. The study will have policy implications for stakeholders involved in the treatment of waste water.

1.6 Organization of the study

The study is organized into five chapters as follows: Chapter one highlights the background of the study, the motivation for the study and the significance of the study among others. Chapter two, the literature review looks at wastewater characteristics, recent application of time series neural network forecasting models and a description of design considerations for short term and long term memory components of time series neural network models, the wastewater treatment data, and the experimental methodology follows in chapter three. In Chapter four the performance of each forecasting model is then evaluated for the prediction of wastewater effluent BOD. Chapter five discusses the model performances and Chapter six gives the conclusion.

CHAPTER TWO

LITERATURE REVIEW

2.1 Wastewater

Municipal wastewater is mainly comprised of water (99.9%) together with relatively small concentrations of suspended and dissolved organic and inorganic solids. Among the organic substances present in sewage are carbohydrates, lignin, fats, soaps, synthetic detergents, proteins and their decomposition products, as well as various natural and synthetic organic chemicals from the process industries. Table 1 shows the levels of the major constituents of strong, medium and weak domestic wastewaters. In arid and semi-arid countries, water use is often fairly low and sewage tends to be very strong, as indicated in Table 2 for Amman, Jordan, where water consumption is 90litres per day per person

Table 1 Major Constituents of Typical Domestic Wastewater

Constituent	**Concentration, mg/l**		
	Strong	**Medium**	**Weak**
Total solids	1200	700	350
Dissolved solids (TDS)[1]	850	500	250
Suspended solids	350	200	100
Nitrogen (as N)	85	40	20
Phosphorus (as P)	20	10	6
Chloride[1]	100	50	30
Alkalinity (as $CaCO_3$)	200	100	50
Grease	150	100	50
BOD_5[2]	300	200	100

Source: UN Department of Technical Cooperation for Development (1985)

[1] The amounts of TDS and chloride should be increased by the concentrations of these constituents in the carriage water.

[2] BOD_5 is the biochemical oxygen demand at 20°C over 5 days and is a measure of the biodegradable organic matter in the wastewater.

Municipal wastewater also contains a variety of inorganic substances from domestic and industrial sources, Table 3 including a number of potentially toxic elements such as arsenic, cadmium, chromium, copper, lead, mercury, zinc, etc. Even if toxic materials are not present in concentrations likely to affect humans, they might well be at phytotoxin levels, which would limit their agricultural use. However, from the point of view of health, a very important consideration in agricultural use of wastewater, the contaminants of greatest concern are the pathogenic micro- and macro-organisms.

Pathogenic viruses, bacteria, protozoa and helminths may be present in raw municipal wastewater at the levels indicated in Table 4 and will survive in the environment for long periods, as summarized in Table 5. Pathogenic bacteria will be present in wastewater at much lower levels than the coliform group of bacteria, which are much easier to identify and enumerate (as $totalcoliforms/100ml$). *Escherichia coli* are the most widely adopted indicator of faecal pollution and they can also be isolated and identified fairly simply, with their numbers usually being given in the form of $faecalcoliforms(FC)/100ml$ of wastewater.

Table 2 Average Composition of Wastewater in Amman, Jordan

Constituent	Concentration mg/l
Dissolved solids (TDS)	1170
Suspended solids	900
Nitrogen (as N)	150
Phosphorus (as P)	25
Alkalinity (as $CaCO_3$)	850
Sulphate (as SO_4)	90
BOD_5	770
COD[1]	1830
TOC[1]	220

Source: Al-Salem (1987)

[1] COD is chemical oxygen demand

[2] TOC is total organic carbon

Table 3 Chemical Composition of Wastewaters in Alexandria and Giza, Egypt

Constituent	Alexandria		Giza	
	Unit	Concentration	Unit	Concentration
EC	dS/m	3.10	dS/m	1.7
pH		7.80		7.1
SAR		9.30		2.8
Na_2^+	me/l	24.60	mg/l	205
Ca_2^+	me/l	1.50	mg/l	128
Mg	me/l	3.20	mg/l	96
K^+	me/l	1.80	mg/l	35
Cl^-	me/l	62.00	mg/l	320
SO_4^{2-}	me/l	35.00	mg/l	138
CO_3	me/l	1.10		
HCO_3^-	me/l	6.60		
NH_4^+	mg/l	2.50		
NO_3	mg/l	10.10		
P	mg/l	8.50		
Mn	mg/l	0.20	mg/l	0.7
Cu	mg/l	1.10	mg/l	0.4
Zn	mg/l	0.80	mg/l	1.4

SAR is sodium absorption ratio

Source: Abdel-Ghaffar *et al.* (1988)

KNUST

Table 4 Possible Levels of Pathogens in Wastewater

Type of pathogen		Possible concentration per litre in municipal wastewater[1]
Viruses:	*Enteroviruses*[2]	5000
Bacteria:	Pathogenic *E. coli*[3]	?
	Salmonella spp.	7000
	Shigella spp.	7000
	Vibrio cholerae	1000
Protozoa:	*Entamoeba histolytica*	4500
Helminths:	*Ascaris Lumbricoides*	600
	Hookworms[4]	32
	Schistosoma mansoni	1
	Taenia saginata	10
	Trichuris trichiura	120

Source: Feachem *et al.* (1983)

KNUST

[?]Uncertain

[1]Based on 100 lpcd of municipal sewage and 90% inactivation of excreted pathogens

[2]Includes polio-, echo- and coxsackieviruses

[3]Includes enterotoxigenic, enteroinvasive and enteropathogenic *E. coli*

[4]*Anglostoma duedenale* and *Necator americanus*

Table 5 Survival of Excreted Pathogens (at 20-30°C)

Type of pathogen	Survival times in days			
	In faeces, night soil and sludge	**In fresh water and sewage**	**In the soil**	**On crops**
Viruses				
Enteroviruses	<100 (<20)	<120 (<50)	<100 (<20)	<60 (<15)*
Bacteria				
Faecal Coliforms	<90 (<50)	<60 (<30)	<70 (<20)	<30 (<15)
Salmonella spp.	<60 (<30)	<60 (<30)	<70 (<20)	<30 (<15)
Shigella spp.	<30 (<10)	<30 (<10)	-	<10 (<5)
Vibrio cholerae	<30 (<5)	<30 (<10)	<20 (<10)	< 5 (<2)
Protozoa	<30 (<15)	<30 (<15)	<20 (<10)	<10 (< 2)
Entamoeba histolytica cysts	<30 (<15)	<30 (<15)	<20 (<10)	<10 (< 2)
Helminths	Many	Many	Many	<60 (<30)
Ascaris lunbricoides eggs	Months	Months	Months	

Source: Feachem *et al.* (1983)

* Figures in brackets show the usual survival time.

2.2. Effects of wastewater usage

2.2.1. Effect on Health

Organic chemicals usually exist in municipal wastewaters at very low concentrations and ingestion over prolonged periods would be necessary to produce detrimental effects on human health. This is not likely to occur with agricultural/aqua cultural use of wastewater, unless cross-connections with potable supplies occur or agricultural workers are not properly instructed, and can normally be ignored. The principal health hazards associated with the chemical constituents of wastewaters, therefore, arise from the contamination of crops or groundwater. Attention has been drawn to the particular concern attached to the cumulative poisons, principally heavy metals, and carcinogens, mainly organic chemicals. World Health Organization guidelines for drinking water quality (WHO, 1984) include limit values for the organic and toxic substances given in Table 6, based on acceptable daily intakes (ADI). These can be adopted directly for groundwater protection purposes but, in view of the possible accumulation of certain toxic elements in plants (for example, cadmium and selenium) the intake of toxic materials through eating the crops irrigated with contaminated wastewater must be carefully assessed.

Table 6 Organic and Inorganic Constituents of Drinking Water of Health Significance

Organic	Inorganic
Aldrin and dieldrin	Arsenic
Benzene	Cadmium
Benzo-a-pyrene	Chromium
Carbon tetrachloride	Cyanide
Chlordane	Fluoride
Chloroform	Lead
2,4 D	Mercury
DDT	Nitrate
1,2 Dichloroethane	Selenium
1,1 Dichlorethylene	
Heptachlor and heptachlor epoxide	
Hexachlorobenzene	
Lindane	
Methoxychlor	
Pentachlorophenol	
Tetrachlorethylene	
2, 4, 6 Trichloroethylene	
Trichlorophenol	

Source: WHO (1984)

Pathogenic organisms give rise to the greatest health concern in agricultural use of wastewaters, yet few epidemological studies have established definitive adverse health impacts attributable to the practice. Shuval *et al*. (1985), reported on one of the earliest evidences connecting agricultural wastewater reuse with the occurrence of disease (Figure 1). It would appear that in areas of the world where helminthic diseases caused by *Ascaris* and *Trichuris* spp. are endemic in the population and where raw untreated sewage is used to irrigate salad crops and/or vegetables eaten uncooked, transmission of these infections is likely to occur through the consumption of such crops. A study in West Germany (reported by Shuval *et al*. 1986) provides additional evidence (Figure 2) to support this hypothesis and further evidence was also provided by Shuval *et al*. (1985; 1986) to show that cholera can be transmitted through the same channel.

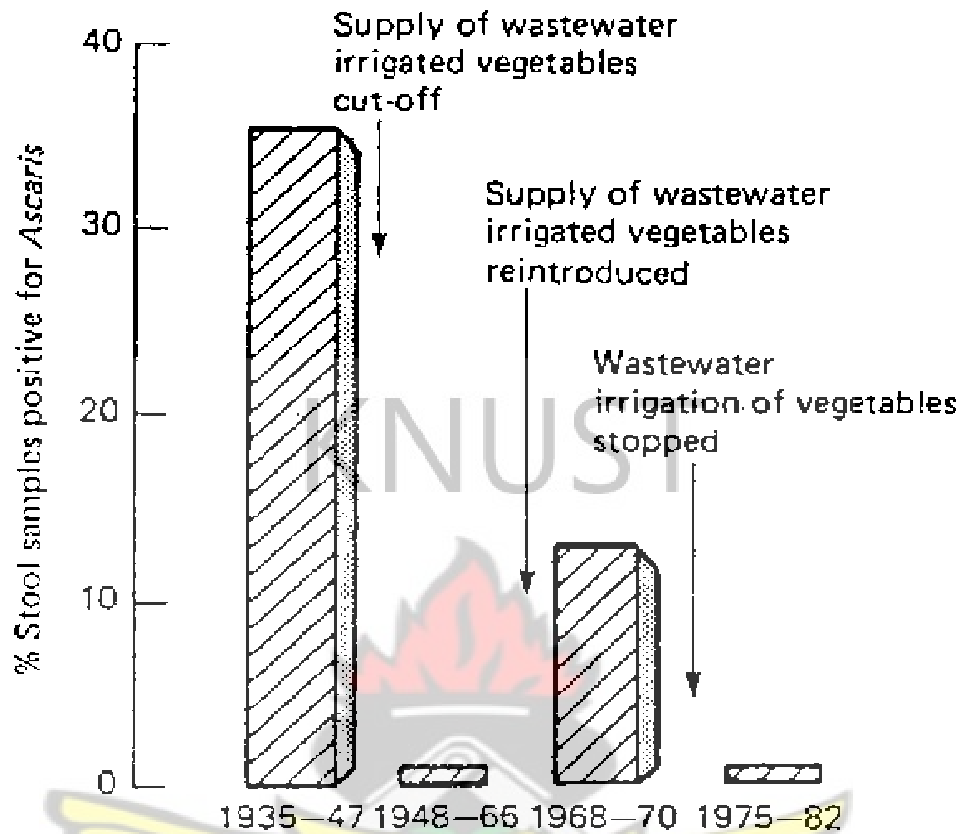


Figure 1 Prevalence of Ascaris-positive stool samples in West Jerusalem population during various periods, with and without supply of vegetables and salad crops irrigated with raw wastewater (Gunnerson, *et al.*1984)

There is only limited evidence indicating that beef tapeworm *(Taeniasaginata)* can be transmitted to the population consuming the meat of cattle grazing on wastewater irrigated fields or fed crops from such fields. However, there is strong evidence from Melbourne, Australia and from Denmark reported by (Shuval *et al.* 1985) that cattle grazing on fields freshly irrigated with raw wastewater, or drinking from raw wastewater canals or ponds, can become heavily infected with the disease (cysticerosis).

Studies in India by Shuval et al. (1986), have shown that sewage farm workers exposed to raw wastewater in areas where *Ancylostoma* (hookworm) and *Ascaris* (nematode) infections are endemic have significantly excess levels of infection with these two parasites compared with other agricultural workers in similar occupations. Furthermore, the studies indicated that the intensity of the Ascaris infections (the number of worms infesting the intestinal tract of an individual) in the sample of sewage farm workers was very much greater than in the control sample. In the case of the hookworm infections, the severity of the health effects was a function of the worm load of individuals, which was found to be related to the degree of exposure and the length of time of exposure to the hookworm larvae. Sewage farm workers are also liable to become infected with cholera if practicing irrigation with raw wastewater derived from an urban area in which a cholera epidemic is in progress (Shuval *et al*. 1985). Morbidity and serological studies on wastewater irrigation workers or wastewater treatment plant workers occupationally exposed to wastewater directly and to wastewater aerosols have not been able to demonstrate excess prevalence of viral diseases.

KNUST

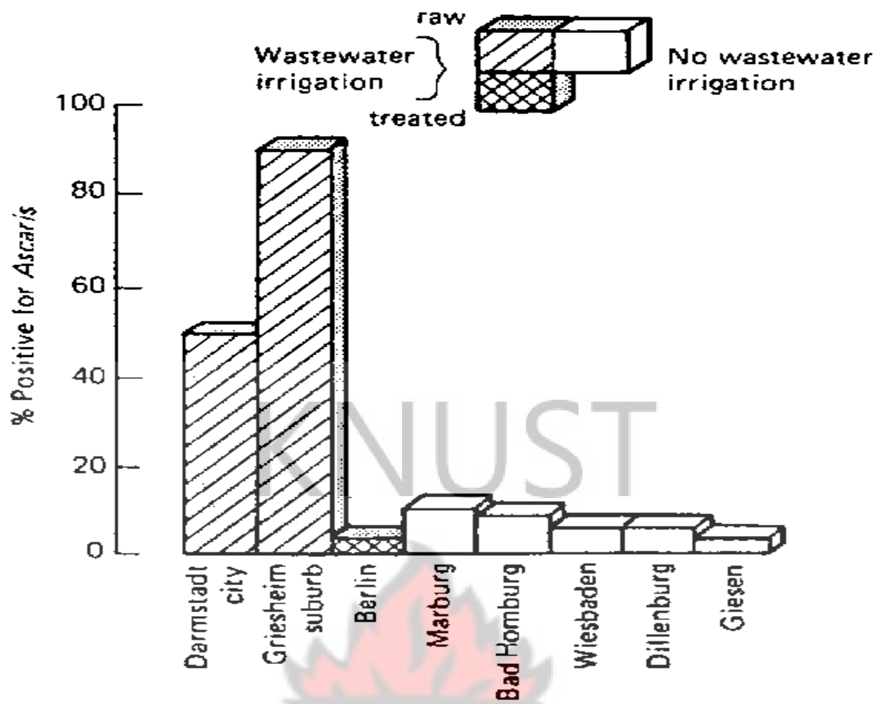


Figure 2 Wastewater irrigation of vegetables and *Ascaris* prevalence in Darmstadt and Berlin, compared with other cities in Germany not practicing wastewater irrigation (Gunnerson, *et al.*, 1984)

No strong evidence has been adduced to suggest that population groups residing near wastewater treatment plants or wastewater irrigation sites are at greater risk from pathogens in aerosolized wastewater resulting from aeration processes or sprinkler irrigation. Shuval *et al*. (1986) suggest that the high levels of immunity against most viruses endemic in the community essentially block environmental transmission by wastewater irrigation.

Finally, in respect of the health impact of use of wastewater in agriculture, Shuval et al. (1986) rank pathogenic agents in the order of priority. They pointed out that negative health effects were only detected in association with the use of raw or poorly-settled wastewater, while inconclusive evidence suggested that appropriate wastewater treatment could provide a high level of health protection.

2.2.2 Effect on agriculture

The quality of irrigation water is of particular importance in arid zones where extremes of temperature and low relative humidity result in high rates of evaporation, with consequent deposition of salt which tends to accumulate in the soil profile. The physical and mechanical properties of the soil, such as dispersion of particles, stability of aggregates, soil structure and permeability, are very sensitive to the type of exchangeable ions present in irrigation water. Thus, when effluent use is being planned, several factors related to soil properties must be taken into consideration. A thorough treatise on the subject prepared by Ayers and Westcot (1985).

Another aspect of agricultural concern is the effect of dissolved solids (TDS) in the irrigation water on the growth of plants. Dissolved salts increase the osmotic potential of soil water and an increase in osmotic pressure of the soil solution increases the amount of energy which plants must expend to take up water from the soil. As a result, respiration is increased and the growth and yield of most plants decline progressively as osmotic pressure increases. Although most plants respond to salinity as a function of the total osmotic potential of soil water, some plants are susceptible to specific ion toxicity.

Many of the ions which are harmless or even beneficial at relatively low concentrations may become toxic to plants at high concentration, either through direct interference with metabolic processes or through indirect effects on other nutrients, which might be rendered inaccessible. Morishita (1985) has reported that irrigation with nitrogen-enriched polluted water can supply a considerable excess of nutrient nitrogen to growing rice plants and can result in a significant yield loss of rice through lodging, failure to ripen and increased susceptibility to pests and diseases as a result of over-luxuriant growth. He further reported that non-polluted soil, having around 0.4 and 0.5 ppm cadmium, may produce about 0.08 ppm Cd in brown rice, while only a little increase up to 0.82, 1.25 or 2.1 ppm of soil Cd has the potential to produce heavily polluted brown rice with 1.0 ppm Cd.

Important agricultural water quality parameters include a number of specific properties of water that are relevant in relation to the yield and quality crops, maintenance of soil productivity and protection of the environment. These parameters mainly consist of certain physical and chemical characteristics of the water.

2.3 Wastewater Analysis

Industrial activities consume a huge amount of natural water, utilizable resources and energy there by discharging enormous wastewater to the natural environment. It is therefore necessary to analyse any industrial wastewater to determine its reuse potential and the degree of treatment required prior to its ultimate disposal or to device suitable measures for the recovery of useful products. It is of great importance in water quality control that the amount of organic matter present in the system be known and that the quantity of oxygen required for its stabilisation be determined. Over the years, different physico-chemical tests have been developed to determine the organic and inorganic content of wastewater (Metcalf and Eddy, 1995). In general, these tests may be divided into those used to measure gross concentrations of organic matter greater than 1mg/l and those used to measure trace concentrations in the range of 10^{-6} to 10^{-3} mg/l. Laboratory methods commonly used today to measure the gross amount of organic matter (greater than 1 mg/l) in wastewater includes the following: Biochemical Oxygen Demand (BOD_5), Chemical Oxygen Demand (COD) and Total Organic Carbon (TOC). These three parameters are used in wastewater treatment operations to estimate the influent and effluent characteristics and treatment efficiency. The use of TOC as an analytical parameter has become more common in recent years especially for the treatment of industrial wastewater. Partly, this is due to the fact that the TOC determinations can be carried out in triplicate within minutes compared with the five days required for the BOD_5 test (Sawyer *et al.*, 1994). Apart from these, the easily measurable parameters for any industrial wastewater include indices like Total Suspended Solids (TSS), Total Dissolved Solids (TDS), Phenol concentration, Ammoniacal Nitrogen (AMN) and Kjeldahl's Nitrogen (KJN), Metcalf and Eddy, (1995).

A review of the existing literature in this field reveals that correlation among these parameters seldom exists. It could be difficult to understand the dynamics of relationship between these parameters because they primarily depend on the process of the target industry, raw material/by-product composition, composition of chemicals discharged in wastewater and thus their nonlinear relationship makes universal generalization difficult.

KNUST

2.4 Measuring BOD_5 Levels

When water samples are collected, dissolved oxygen (DO) and the PH are measured at the time of collection. Samples are then placed in bottles full to the brim and sealed off by a lid. The samples are covered completely with Aluminum foil and placed in a dark place. This limits the photosynthesis, which could happen with captured algae. After five days the bottles are uncorked and the dissolved oxygen is probed. The difference between the first dissolved oxygen and last dissolved oxygen of the sample is called the biochemical oxygen demand (BOD). A low number generally means little pollution and/or little aerobic activity. A high number means the opposite.

2.5 Neural networks and their applications

2.5.1 Introduction

One of the most popular data-driven techniques attributed by various authors to machine learning, data mining, soft computing etc. is an Artificial Neural Network (ANN). An ANN is an information processing system that roughly replicates the behaviour of a human brain by emulating the operations and connectivity of biological neurons (Tsoukalas and Uhrig, 1997).It performs a human-like reasoning, learns the attitude and stores the relationship of the processes on the basis of a representative data set that already exists.

Depending on the structure of the network, usually a series of connecting neuron *weights* are adjusted in order to fit a series of inputs to another series of known outputs. When the weight of a particular neuron is updated it is said that the neuron is *learning*. The training is the process that neural network learns. Once the training is performed the verification is very fast. Since the connecting weights are not related to some physical identities, the approach is considered as a black-box model. The adaptability, reliability and robustness of an ANN depend upon the source, range, quantity and quality of the data set.

During the last decade, ANNs evolved from only a research tool into a tool that is applied to many real world problems: physical system control, engineering problems, statistics, even medical and biological fields. The number of European patents obtained in the last decade corroborates the trend of increased applications of ANNs (Kappen, 1996).

2.5.2 Basic elements in neural network structure

As discussed in 2, the ANN performs fundamentally like a human brain. The cell body in the human neuron receives incoming impulses via *dendrites* (receiver) by means of chemical processes. If the number of incoming impulses exceeds certain threshold value the neuron will discharge it off to other neurons through its *synapses*, which determines the impulse frequency to be fired off (Beale and Jackson, 1990).Therefore, processing units or neurons of an ANN consists of three main components; synaptic weights connecting the nodes, the summation function within the node and the transfer function (see Figure 4). *Synaptic weights* characterise themselves with their strength (value) which corresponds to the importance of the information coming from each neuron. In other words, the information is encoded in these strength-weights. The *summation function* is used to calculate a total input signal by multiplying their synaptic weights and summing up all the products.

2.5.3 Neural network structures

Structure of an ANN can be classified into 3 groups as per by the arrangement of neurons andthe connection patterns of the layers: feed forward (error back propagation networks), feedback (recurrent neural networks and adaptive resonance memories), self-organizing (Kohonen networks). Also neural networks can be roughly categorized into two types interms of their learning features: *supervised* learning algorithms, where networks learn to fitknown inputs to known outputs, and *unsupervised* learning algorithms, where no desiredoutput to a set of input is defined. The classification is not unique and different researchgroups make different classifications. One of the possible classifications is shown in Figure 3

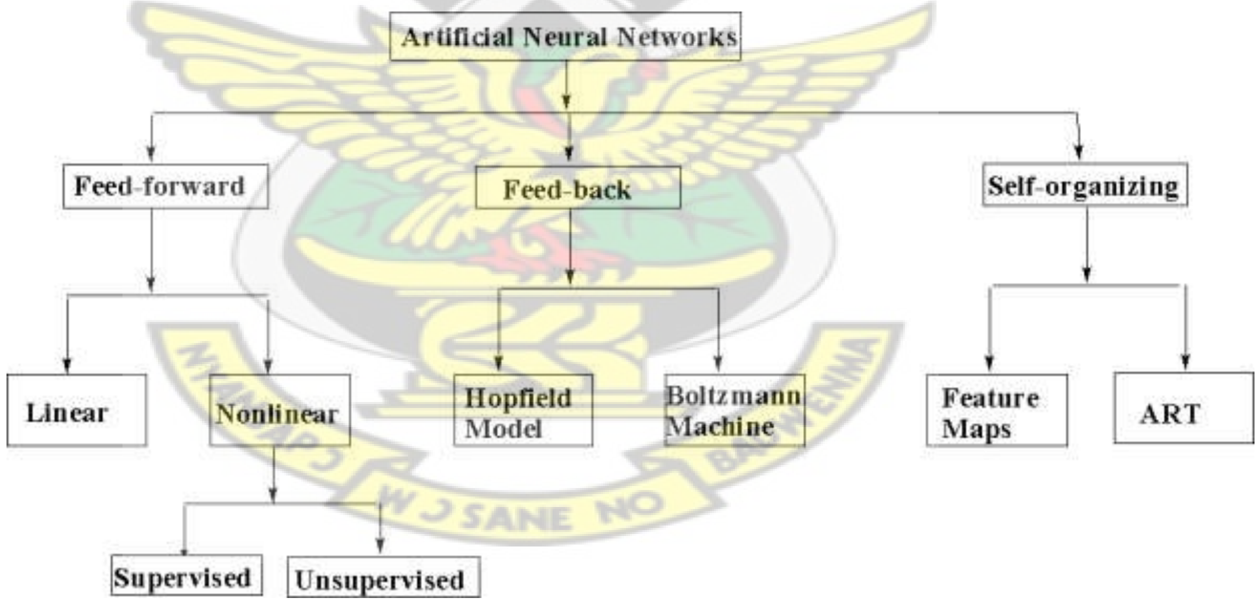


Figure 3 Neural network classifications

The feed forward neural networks consist of three or more layers of nodes: one input layer, one output layer and one or more hidden layers. The input vector *x* passed to the network is directly passed to the node activation output of input layer without any computation. One or more hidden layers of nodes between input and output layer provide additional computations. Then the output layer generates the mapping output vector *z*. Each of the hidden and output layer has a set of connections, with a corresponding strength-weight, between itself and each node of preceding layer. Such structure of a network is called a *Multi-Layer Perceptron*(MLP). Figure 4 shows a typical multi-layer perceptron.

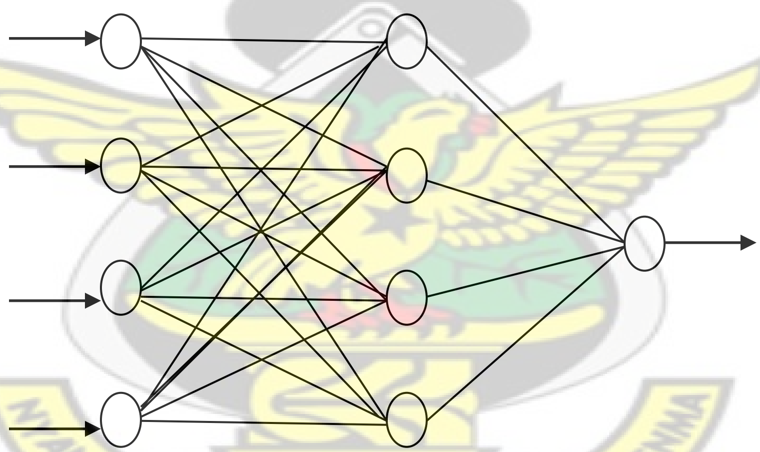


Figure 4 A fully connected multi-layer perceptron

KNUST

The feedback neural networks have loops that feedback information in the hidden layers. In Self-Organising Feature Maps (SOFM) the multidimensional input space is mapped into two or three dimensional maps by preserving the necessary features to be extracted or classified. An SOFM consists of an input layer and an output map.

2.5.4 Neural Networks for Time-Series Forecasting

Forecasting future events is an important task in many practical situations. These situations may include the forecasting of river inflow; temperature, rainfall or weather; electricity load or energy; economic, stock market or currency exchange rate; and numerous other scenarios in which future events need to be predicted based on available current and past information. In general, this involves a series of event outcomes that are related or correlated, and correlation dynamics are used to predict the next one or several outcomes in the series. Many cases involve a time-series consisting of the observations of an event over a period of time, and require a prediction of the event for a particular time (hour, day, month), based on the outcome of the same event for some past instances of time. Time-series forecasting is an important area of forecasting in which past observations are analyzed to develop a model describing the underlying relationship in a sequence of event outcomes. The model is then used to forecast future events that have not yet been observed

A time-series can be either linear or nonlinear. In a linear time-series, the next outcome is linearly related to the current outcome. In contrast, in a nonlinear time-series, the next outcome has a nonlinear relationship to the current outcome. A main feature of any time-series is autocorrelation, which is simply the correlation of the value of the observation at a particular time with observations at previous instances in time. In general, the autocorrelation is higher for events in the immediate past and decreases for observations in the distant past. The numbers of steps that are significantly related are known as lags, and for a time-series, these steps are called time lags. If only the immediate past event has a significant relation to the next observation, time lag is one, and so on. Depending on the problem, the related time lags can all be past observations from the current time to a specific time in the past, or they can be some intermittent past observations between a current and past instance of time. It can be intuitively expected that the effects of the past outcomes are captured in the current observation of a variable, and that therefore, the current observation should be a good predictor of future observations. However, practice, it is known that time lags improve forecasting accuracy.

2.5.5 Neural Networks for Nonlinear Time-Series forecasting.

The power of neural networks lies in their ability to capture nonlinear relationships inherent in data. Whereas linear models depict a linear relationship between the current and next observations, neural networks portray a nonlinear relationship between the two. This can be described as

$y_t = f(y_{t-1}, y_{t-2}, \dots y_{t-p}) + \varepsilon_t$ …………………………………………..2.1

Where $f(y_{t-1}, y_{t-2}, \dots, y_{t-p})$ is the nonlinear function that maps a series of past observations nonlinearly to the next outcome. This function is the neural network model. The last component in Equation 2.1 is the error, which is expected to be a random variable with a mean of zero and variance σ^2

2.5.6 Types of Neural Network for Nonlinear Time-Series Forecasting

Two types of neural network models that have been successfully used for various time-series forecasting tasks are focused time-lagged feed forward networks and dynamically-driven recurrent (feedback) networks. Both are based on the feed forward, multilayer networks. A schematic diagram of the focused time-lagged feed forward network is presented in Figure 5, in which the time is embedded externally using a short-term memory filter that acts as input to a static feed forward network such as Multi-Layer Perceptron (MLP) (Haykin, 1991). The short-term memory filter basically incorporates time lags. Thus a short-term history of the time-series is presented as inputs to a static network.

The concept of a recurrent network is entirely different from that of a focused time-lagged network. It attempts to incrementally build the autocorrelation structure of a series into the model internally, using feedback connections relying solely on the current values of the input(s) provided externally (Haykin, 1991) The idea behind such networks is that a network should learn the dynamics of the series over time from the present state of the series, which is continuously fed into it, and that the network should then use this memory when forecasting.

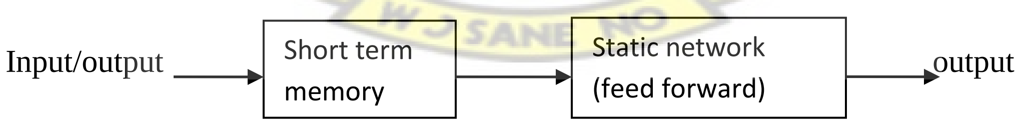


Figure 5 focused time –lagged (short-term memory) feed forward network

2.5.6.1 Focused Time Lagged Feed Forward Networks

In focused time lagged (or time lagged) networks, input variables presented to a feed forward network, such as an MLP, at time t are fed again with a unit time delay at next time step. This is repeated recursively up to the required lag length. In this manner a short term memory of autocorrelation structure is built into the network by externally feeding inputs as lags. Thus, in time lagged network, the model sought is of the form depicted in Equation 2.1

KNUST

2.5.6.2 Spatio -Temporal Time Lagged Network

In some situations, the response of a time- series is affected by one or more time- series. Accurate forecasting of such situation required the incorporation of memory dynamics of each of the time- series. For example, the energy requirement of a building on a particular day is correlated with that of the previous time steps, as well as the temperature over the past several days. A spatio temporal time lagged networks implements this by filtering short term memory response for each spatial dimension. Such a neuron is depicted in Figure 6

NYANSAPɔ WɔSANE NO BADWENMA

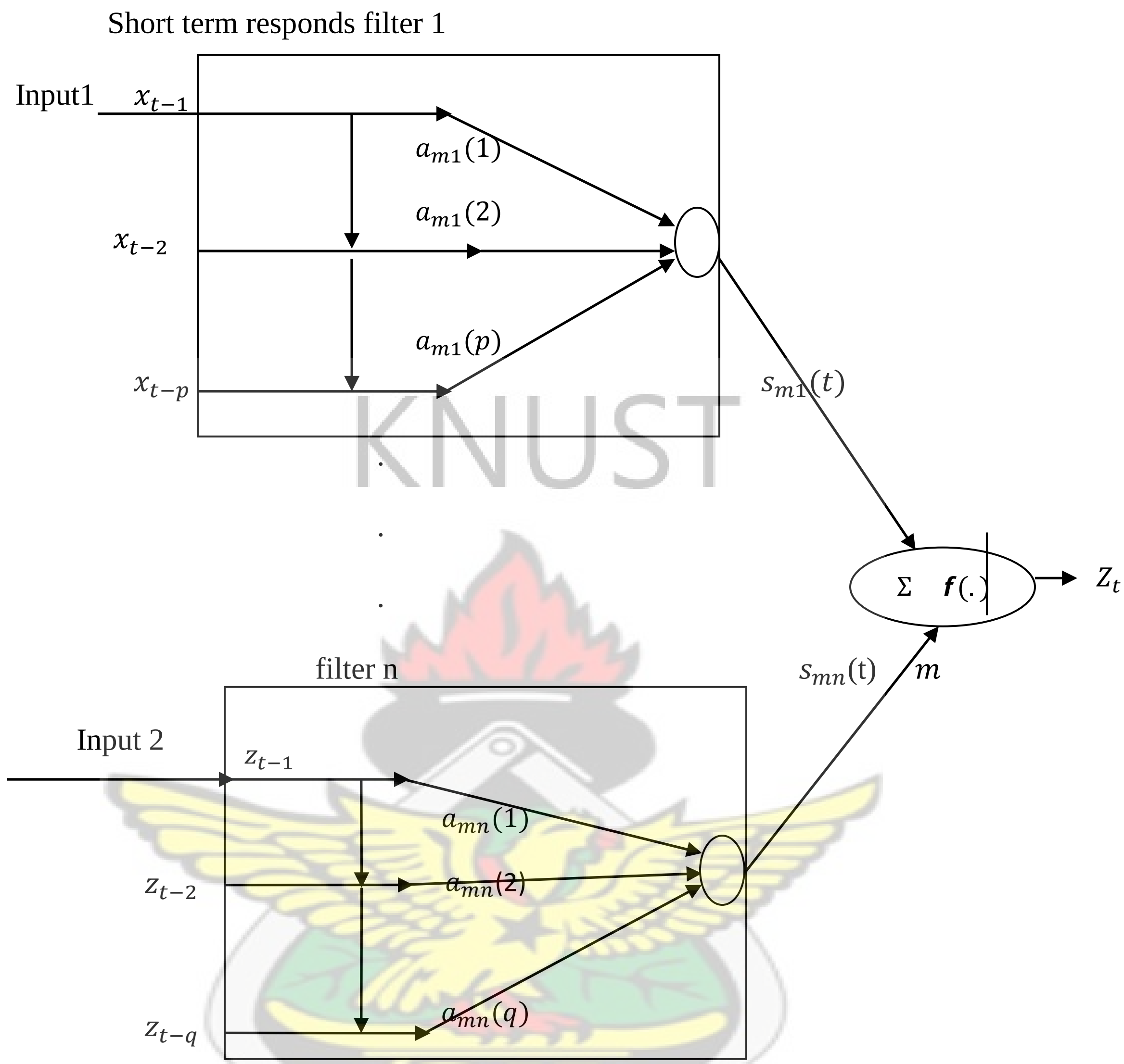


Figure 6 Multi input time lagged neuron(or spatio-temporal filter)

The neuron in figure 6 receives time delay inputs of various orders (lag length) from several inputs. In time- series one of these inputs variables is to be forecast. The other variables are those that influence forecast accuracy. The neuron m receives filtered short-term memory of several variables via the corresponding weights a_{mi}. The short term memory weight of one variable is embedded in the weighted sum of the time lag $s_{mi}(t)$

The weighted sum for input one is

$s_{m1}(t) = \sum_{l=1}^{p} a_{m1}(l)\, x_{t-1}$ ……………………………..2.2

Where $a_{m1}(t)$ is the weight from input one with lag l to neuron m and x_{t-1} is the input corresponding to lag l extending from l to p, the order of the short term memory. The total weight vector a_{m1} and the input vector x of input one are

$a_{m1=}[a_{m1}(1), a_{m1}(2), \dots, a_m(p)]$………………………… …2.3

$x = [x_{t-1}, x_{t-2}, \dots, x_{t-p}]$…………………………………… 2.4

2.6 Time-Delay Neural Network

One of the most popular techniques for dealing with the modeling of dynamic systems is the time-delay neural network (TDNN), which was first introduced by Lang and Hinton (1988) and Waibel *et al*. (1989), and was popularized by Haykin,S., (1994). It is a multilayer feed forward network whose hidden neurons and output neurons are replicated across time. It should be pointed out that the TDNN topology is in fact embodied in a Multilayer perceptron (MLP) in which each synapse is represented by a finite-duration impulse response (FIR) filter.

In the standard MLP, the contribution of a synapse to the linked neuron is the product of the corresponding input and weight of the synapse, which is expressed as

$S_i = x_i w_i$……………………………………………………………2.5

In a TDNN, the input is a time series, denoted as $x_i(t)$and the weight of the synapse is replaced by an FIR filter whose impulse response is denoted as $w_i(t)$

The TDNN used in this study are feed forward multilayer perceptrons, where the internal weights are replaced by finite impulse response (FIR) filters (Figure 7). By using FIR filter as synaptic connection, each weight is replaced by a weight vector

$a_{mi}=(a_{mi}(1), a_{mi(2),\dots,}\ a_{mi})$……………………………………………………..2.6

Where a_{mi} is the vector carrying signals from neuron i to neuron m. The output state of neuron i is a p dimensional vector $x_i(t)$, storing the history of the state of neuron i through $p$previous time steps.

$x_i(t) = (x_i(t-1), x_i(t-2)\dots x_i(t-p))$………………2.7

The input signal to every neuron at the application of the input pattern *t* includes the output signals from the previous layer at time steps $t, t-1, t-2, \dots, t-p$ where p is the delay associated with the corresponding weight vector.

KNUST

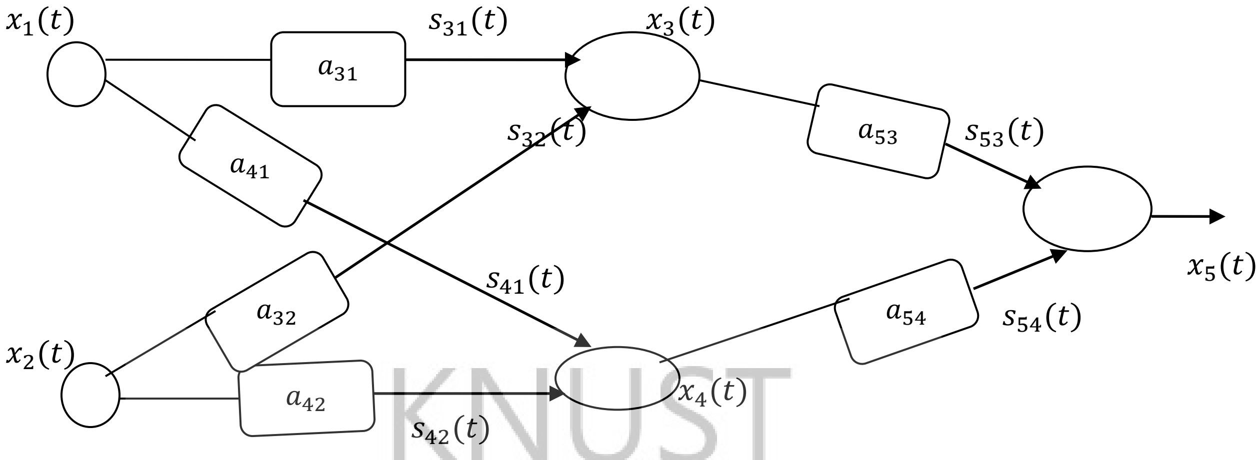


Figure 7(a) Three neuron with FIR filter ($\boldsymbol{a_{mi}}$) as synaptic connection

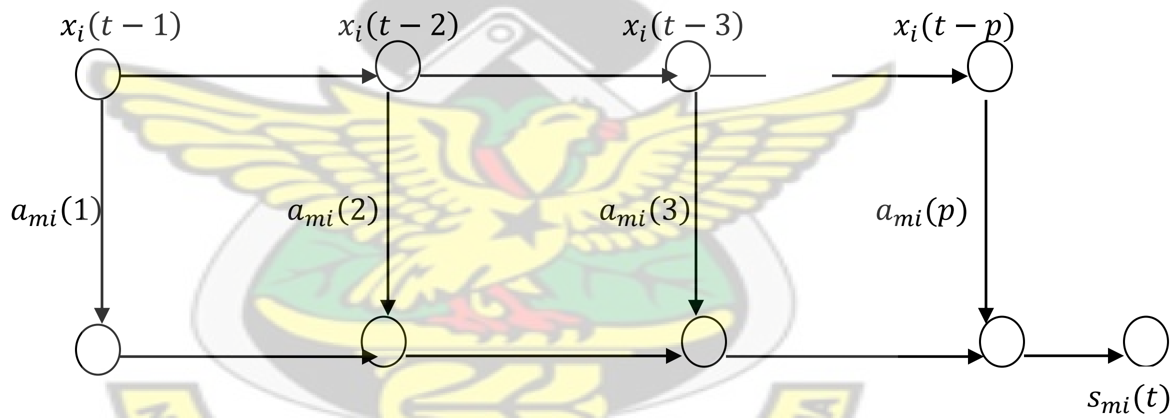


Figure 7 (b) Expanded view of FIR synaptic connections of TDNN. FIR filters build internal memory into the network

Taking the inner product of $x_i(t)$ and a_{mi} , we get the output of the weight filter as

$s_{mi(t)=}a_{mi}x_i(t)$……………………………………2.8a

$=\sum_{l=1}^{p} a_{mi}\,(l)x_i(t-1)$…………………………2.8b

Where p is the length of the impulse response series. From Equation. (2.8b), it can be seen that in a TDNN, the contribution of a synapse relates not only to the input at time t, but also the data at times $t-1,\ t-2,\dots,t-p$. The weights in the standard MLP are split into a weight series of length m. That is in the TDNN, the time series is represented by the current and $(p-1)$ lagged values.

KNUST

Other input can have different order or lag length and therefore can have a different number of inputs and weights.

When all the sum are calculated for n inputs variables, there are n weighted sum, $s_{m1}(t)$,$s_{m2}(t,\dots,s_{mn}(t))$. The neuron m simply sums these to get the total weighted sum,$s_m(t)$ of the short-term responses for time t, for the neuron as

$s_m(t) = \sum_{i=1}^{n} s_{mi}(t)$…………………………………2.9

The activation potential,$s_m(t)$ of the neuron $m$is processed nonlinearly by the nonlinear activation function of the neuron to produce an output, $y_m(t)$ ie

$y_m(t) = f(s_m(t))$…………………………………….....2.10

WJ SANE NO BADWENIMA

For all neurons, we used the hyperbolic tangent activation function which usually accelerates training (Haykin, 1998).

$f(x) = (1-e^{-x})/(1+e^{-x})$……………..……………2.11

2.7 Back Propagation Training Algorithm

One of the most commonly used techniques for learning in neural networks is called

Back propagation. In order for the weights of the neural network connections to be adjusted, an error first needs to be calculated between the predicted response and the actual response. The following formula is commonly used for the output layer:

$Error_i \ = \ Output_i(1 - Output_i)(Actual_i - Output_{i)}$………..2.12

Where $Error_i$ is the error resulting from$neuron_i$, $Output_i$ is the predicted response value and $Astual_i$ is the actual response value.

KNUST

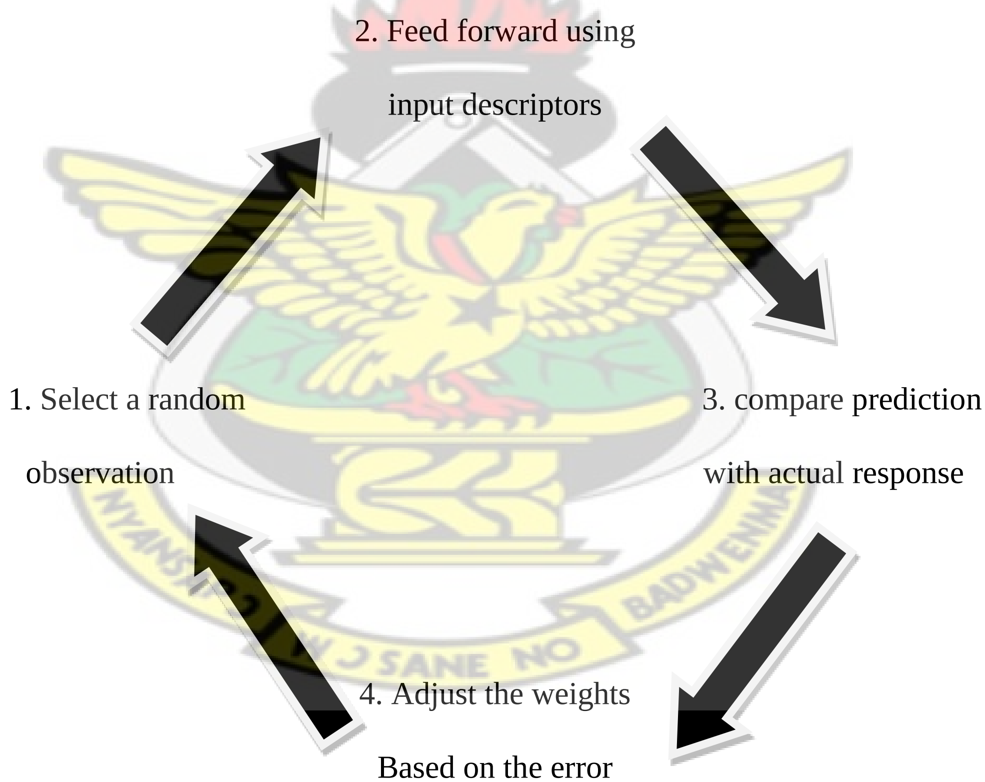


Figure 8 Learning process in neural networks

Once the error has been calculated for the output layer, it can now be back propagated, that is, the error can be passed back through the neural network. To calculate an error value for the hidden layers, the following calculation is commonly used:

$Error_i = Output_i(1 - Output_i)\sum_{j=1}^{n} Error_j\, w_{ij}$……..2.13

Where $Error_i$ is the error resulting from the hidden neuron, $Output_i$ is the value of the output from the hidden neuron, $Error_j$ is the error already calculated for the jth neuron connected to the output and w_{ij} is the weight on this connection.An error should be calculated for all output and hidden layer neuron. Errors for hidden layer neuron use errors from the neuron their output is attached to, which have already been calculated. Once the error has been propagated throughout the neural network, the error values can be used to adjust the weights of the connections using the formula:

$w_{ij} = w_{ij} + l \times Error_j \times Output_i$…………………2.14

Where w_{ij} is the weight of the connection between neuron i and j, $Error_j$ is the calculated error for neuron j, $Output_i$ is the computed output from node i, and l is the predefined learning rate. This takes a value between 0 and 1. The smaller the learning rate value, the slower the learning process. Often a learning rate is set high initially and then reduced as the network fine-tunes the weights.

2.8 Applications of Nonlinear Neural Network Models in Complex Forecasting Domains

Coulibaly *et al.*, (2000) used a feed forward neural network to forecast daily inflow to a Northern Quebec reservoir, a multivariate hydrological time series. Their results suggest that the neural network is effective for improved prediction accuracy of water inflow.

The observation that short term electricity demand exhibits nonlinear behavior has motivated several authors to investigate neural network forecasting models (Darbellay and Slama, 2000; Hobbs *et al.*, 1998). Jorquera *et al*. (1998), studied neural networks to predict the maximum ozone levels in Santiago, Chile. In a similar study, Prybutok *et al.*, (2000) concluded that neural networks provide better forecasts of the maximum ozone concentration in Houston, Texas. The forecast of rainfall amounts is also extremely difficult, involving both temporal and spatial outcomes that are dependent on many variables and complex atmospheric processes. Luk *et al.*, (2001), forecast one time-step rainfall using recurrent and time delay neural networks and found that these models produce reasonable predictions.

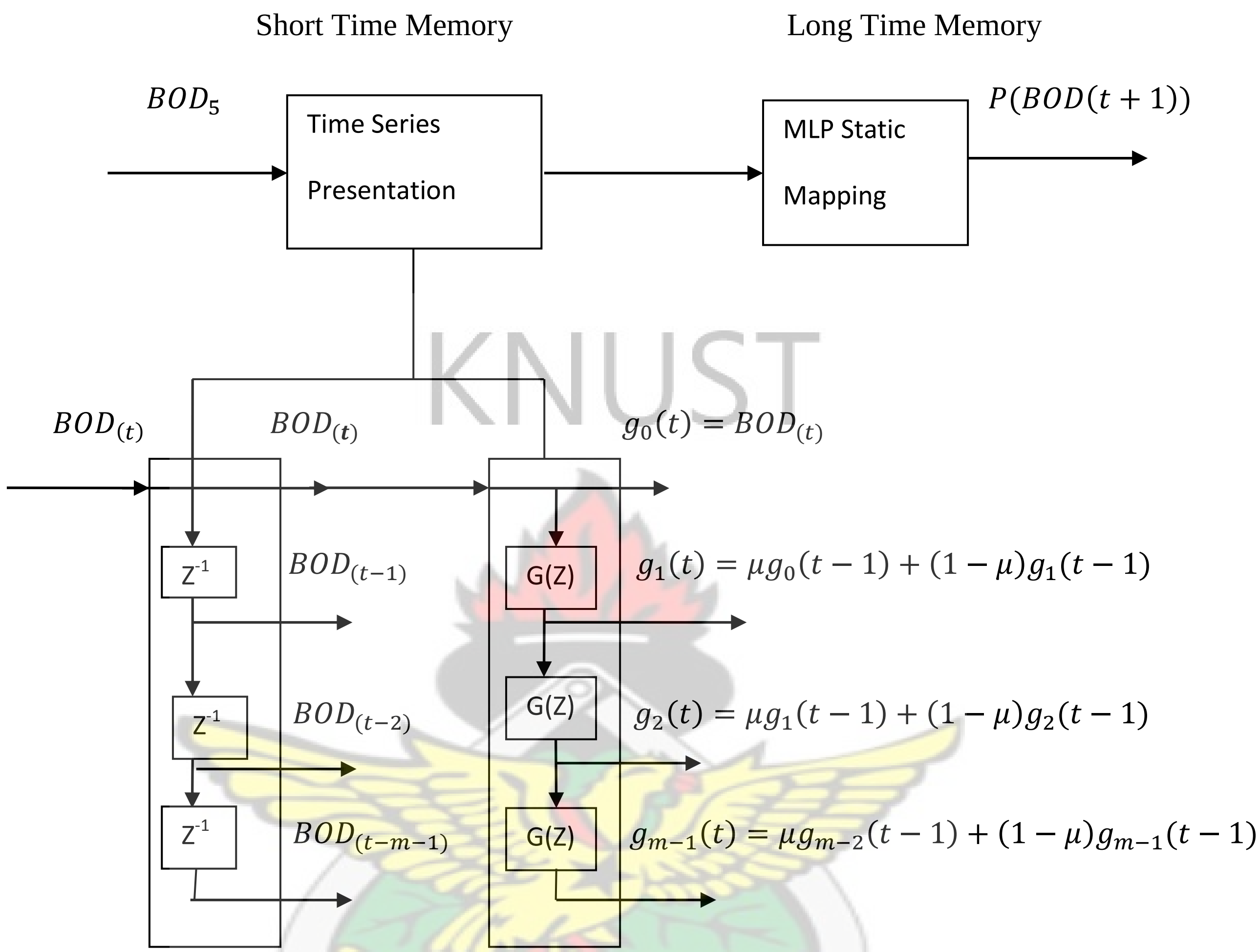
2.9 Application of Nonlinear Neural Network Models on Wastewater

There have been several attempts to model wastewater treatment applications with neural networks. Zhu *et al.* (1998), used a time delay neural network (TDNN) to predict BOD and several other process measurements in the effluent of a wastewater treatment plant, but they do not report any forecasting performance metrics or comparative analysis. Wen and Vassiliadis (1998), integrated a neural network into a hybrid wastewater treatment model to study the large temporal variations that occur in wastewater composition, concentration, and flow rate. Boger (1997), employed a neural network model to predict ammonia levels in the effluent of a wastewater treatment process. Gontarski *et al.* (2000), developed a neural network simulation model for a wastewater treatment plant and concluded that an artificial neural network can be used to establish improved operating conditions.

Since the data for the wastewater treatment process is nonlinear and dynamic (i.e., the statistical properties change over time), a network with a short term memory is required for modeling the temporal patterns in the data. The traditional multilayer perceptron neural network (MLP) can only learn a static mapping of input variables to a forecast value. This is of limited utility for forecasting wastewater BOD since the MLP alone cannot model the temporal or dynamic aspects of the wastewater system where the transport (retention time in system) is not a constant, but changes with flow rate of water and drifting of the process to new operating conditions or states. Two major subcomponents (see Figure 9) are necessary for the development of a neural network time series model, *short term memory*, for representing the time series structure of the data, and *long term memory*, the traditional static mapping from the time series representation to a predicted output value.

2.9.1 Short Term Memory

The purpose of the short term memory is to capture the time series structure of the data in a low dimensional space that preserves the dynamic properties of the system. Two important properties of short term memory are the memory depth, m, (i.e., how many data values are necessary) and memory resolution (i.e., the precision of the representation of these data values). The selection of the memory depth, m, is an important determinant of the neural network forecasting accuracy. A large value for the memory depth increases the dimensionality of the subsequent MLP neural network and the number of network weights to be optimized during network training.



Memory by Delay (TDNN) Memory by Feedback (GMNN)

Figure 9 Neural Network Time Series Model

A small memory depth risks the loss of critical information necessary to define the system dynamics. Figure 9 includes two alternative strategies for representing the time series structure, *memory by delay* and *memory by feedback*. The TDNN architecture uses *memory by delay* for the short term memory representation. The input to the TDNN memory by delay structure is the current time series value (in this application BOD (t)); the output is a window of time-lagged input values of fixed size ,*m*. This output includes the current time series observation at time t and values for the previous $m - 1$ time periods. This set of time delayed observations can be implemented by either partitioning the input data into sequences of length m, or including the delay mechanism in the network input layer. There is no free parameter in the memory by delay structure, so traditional back-propagation cannot be used to adapt the TDNN memory structure to the input data.

KNUST

The Gamma memory implements a *memory by feedback* (De Vries and Principe, 1992). This is detailed in Figure 10 The advantage of Gamma memory is that the parameter μ can be adapted jointly with the long term memory weights during network training.

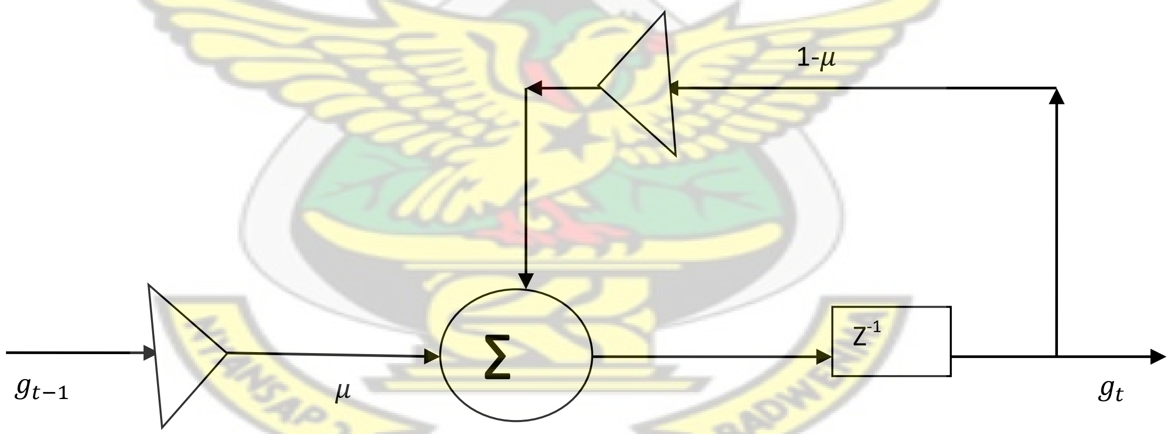


Figure 10 Definition of G (Z)

The adaptation process means that the memory depth and memory resolution will be determined by the data as opposed to the more arbitrary decision required for the TDNN memory implementation. Adaptation of the short term memory requires a modification to the traditional back propagation of error method to account for the fact that the error surface has time gradients. The resulting algorithm is referred to as back propagation through time (Haykin, 1994).

2.9.2. Long Term Memory

The outputs of the short term memory are propagated to an MLP hidden layer and output layer to determine a predicted BOD level. The design of the long term MLP structure involves familiar issues in neural modeling that focus primarily on the configuration of the hidden layer. This involves determining the number of hidden layers, the number of neurons for each hidden layer, and the form of the nonlinear transformation (i.e., sigmoid for multilayer perceptron or Gaussian for radial basis function).

2.10. Important Network Parameters

Good network architecture requires selecting the most dependable values of network parameters like: number of hidden layers, the number of neurons in the hidden layer NH, the activation function $f(x)$, the learning rate of the network η, epoch size ε, momentum term α and training cycles TC. The best values for parameters: η, α, NH, and TC are normally estimated by a trial and error approach. The learning rate η and momentum α can play an important role in the convergence of the network. The η value of a network affects the size of steps taken in weight space (Maier and Dandy, 1998). If η is too small, the algorithm would take more time to converge. The momentum term α accelerates the convergence of the error during the learning process by adding a fraction to the precious weight update. The values of η and α vary between 0−1 and are normally estimated by trial and error (Hamed *et al.*, 2004).

KNUST

CHAPTER THREE

MATERIALS AND METHODS

3.1. Introduction

A neural network is a mathematical model that makes predictions based on a series of input descriptor variables. Like all prediction models, it uses a training set of examples to generate the model. This training set is used to generalize the relationships between the input descriptor variables and the output response variables. Once a neural network has been created, it can then be used to make predictions

3.2. Data Collection

The data used in this study was taken from the source: U.S.A. EPA (1973), Monitoring Industrial Wastewater, Washington, D.C. There are 99 observations, each measured on a - 4 hour composite sample, giving six observations daily for sixteen days, plus three observations on the seventeenth day. The survey was undertaken to estimate the average BOD and to estimate the concentration. The information is needed to design a treatment process. The BOD_5 data for the seventeen days were used as inputs and output variables in this study.This data was used against the backdrop that such data was not readily available or could not be found on an industrial wastewater database here in Ghana but answers to such request were negative. An extensive desk study and internet search was therefore carried out to gather BOD_5 data on industrial wastewater. However, it was hoped that when we have been able to establish and measure such data the model would still be applicable.

The BOD values for daily measurements are plotted in Figure 3.1. It is obvious that this data has a time series structure and periods of significant disturbances to the process.

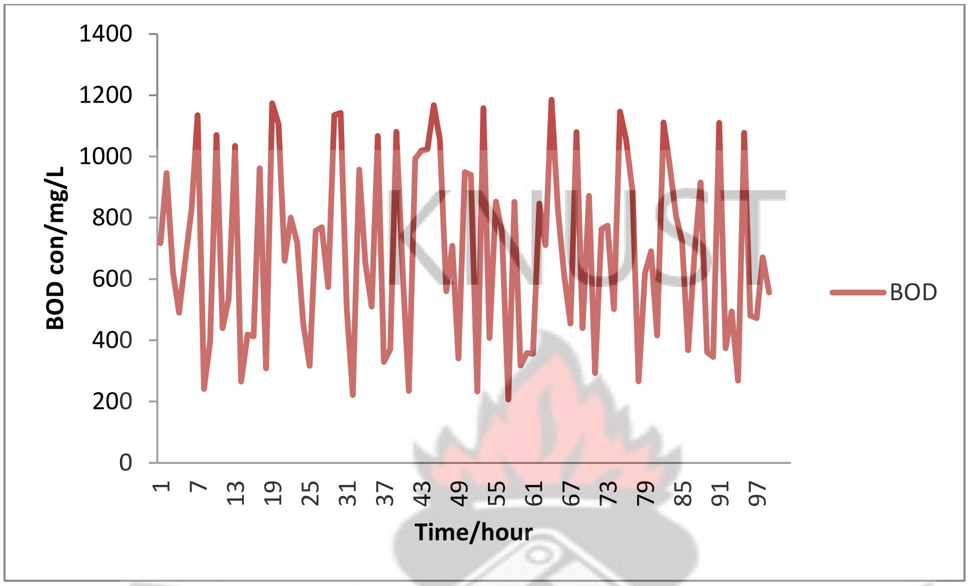


Figure 11 Wastewater BOD concentration with time series

3.3 Research Instruments

TDNN is the artificial neural network used in this study. TDNN as artificial neural networks (ANNs) has a computing system similar to human nervous system where several neurons are connected and communicate with each other to arrive at a decision. The most important advantage of ANNs and for that of TDNN over other modeling techniques is their ability to model complex and nonlinear processes without any a priori knowledge of input and output relationships. A time- delay neural network (TDNN) model is proposed and implemented in this project to predict BOD level of wastewater. Figure 12 gives the structure of the TDNN model of the wastewater treatment system.

The inputs of the model are divided into two groups according to the day when the inputs were measured.

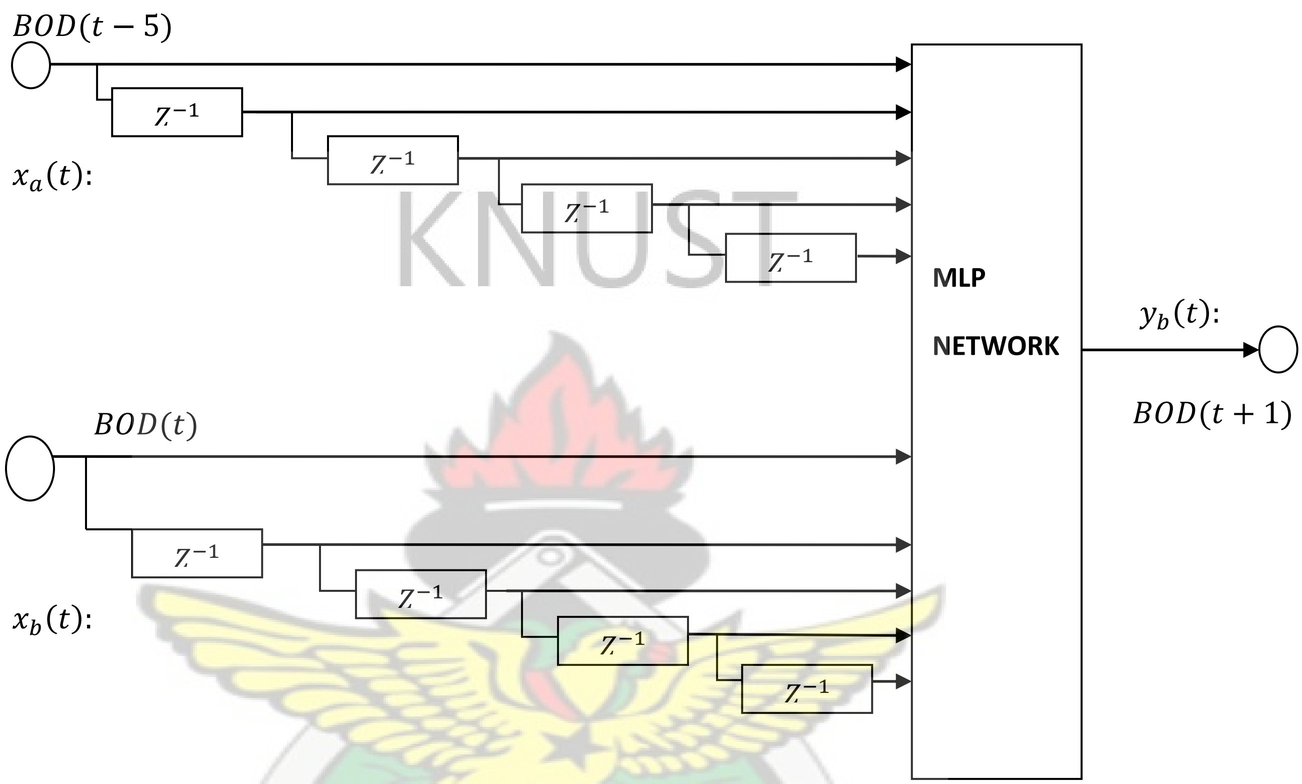


Figure 12 Input/output structure of the TDNN model of the proposed wastewater treatment system.

Let $x_a(t)$, and $x_b(t)$be the input vectors on day a and b and $y_b(t)$be the output scalar on day b. Considering a $5-day$ delay from day a to day b and choosing $m = 4$, the TDNN wastewater treatment system model is expressed as

$y_b(t) = f(x_b(t), x_b(t-1), x_b(t-2), x_b(t-3), x_b(t-4), x_a(t-5), x_a(t-6), x_a(t-7), x_a(t-8), x_a(t-9)$……………………………………………3.15

A training set from the same historical data that is used to train the standard MLP network, is also used to train TDNN model using the back propagation training algorithm.

In developing the model for the estimation of BOD_{t+1}(output parameter) from input parameters (BOD_0,BOD_1,BOD_2,BOD_3,BOD_4,BOD_5,BOD_6,BOD_7,BOD_8, BOD_9,BOD_{10}), a relationship between input and output parameters is developed by a training process, which is explained in the next paragraph. The information flows from the input layer towards the output layer through the hidden layer where the input layer receives information from the input data, processes it in hidden layer(s), and produces the output. A total effect, i.e. the sum of all effects of all the inputs on a given node, is determined by adding the product of each input and the corresponding weight, and then the total effect is transformed to the output using an activation function, and finally the output is transferred to each neuron of the next layer through a connection weight. A nonlinear transfer function as an activation function is used to process the sum and to generate the results. The sigmoid function is one of the transfer functions most commonly used.

3.4. Research Design

A TDNN must be trained before testing its ability to predict the response variable. The aim of learning is to determine the set of weights that minimize the error function. A back-propagation method is used in the training of the TDNN to adjust the connection weights and to obtain the best match between the TDNN's response and the expected response. In a training procedure, firstly, the weights are initialized with a set of random values. Then the weights are systematically updated based on the training rule. After several attempts, the training process is terminated when the difference between the measured and predicted values is less than the specified criteria. One of the common problems arising from the training process is the over-training or over-fitting which might produce incorrect results. The over-training of TDNN depends on the use of the optimal number of nodes in the hidden layer. If the number of hidden nodes is too small, then the network may have insufficient degrees of freedom to learn the process adequately. If the number is too high, then the training takes a long time and the network may over-fit the data. The number of nodes in the hidden layers is determined by a trial-and-error procedure in this study.

In the application of TDNN technique, the parameters of each forecast model were first estimated from a series of consecutive daily BOD observations. The series used to estimate model parameters are referred to as a data window. The size of the data window selected for this research consists of 60 observations, representing approximately 10 days of historical data. Once the parameters are estimated, the forecast models generate one-step forecasts. These forecasts are predicted values of the BOD level for the daily BOD observation immediately following the data window. This is an independent hold out prediction since this BOD value was not used to estimate model parameters. For instance, a data window made up of the first 60 observations (one through 60) is used to forecast BOD one day into the future (i.e., day 11). The data window moves iteratively through the data generating 5000 cycles of estimation and forecast. A single estimation-forecast activity is referred to as a forecast cycle.

3.5 Data Analysis

A total of six (6) TDNN based models were evaluated in this study for predicting the BOD of wastewater. These models are shown in Table 7 The lengths of short term delays, m, are varied from 2 to 12 in increments of two and the number of hidden neurons varied from 2 to 14 by increments of four. The error measures for all TDNN configurations are given in Tables and the MSE results are plotted in Figures. All simulations were carried out by the neural network toolbox of the software NEURALSOLUTIONS. The models performance was evaluated by the mean square error values for training obtained directly from the software, while the test data was evaluated using mean average percentage error. Low MSE and low MAPE values theoretically mean that the predictions are precise and accurate.

CHAPTER FOUR

RESULTS

4.1Prediction of BOD

The Time Delayed Neural Network (TDNN) is used in this study to predict BOD levels, varying the number of time delays, m , from 2 to 12 in increments of two and the number of hidden neurons varied from 2 to 14 by increments of four. The two error measures for all TDNN configurations are given in Table 7 and the MSE results are plotted in Figure 13. The MSE performance indicates that lower error levels are achieved with a fairly large number of neurons in the hidden layer. Series 4 indicates lower error resulted from the MSE of each of the six models when their number of neurons was 14.

Table 7 Comparison of the of Various TDNN Models and Their Best MSE and MAPE

Model	Sub model	m	Input parameters(time series)	Output	MSE	MAPE	Training Count	LR
MO1	$MO1_d$	2	BOD_0, BOD_{-1},	BOD_1	10.0088	17.7154	1000	0.2
MO2	$MO2_d$	4	BOD_0,BOD_{-1}, BOD_{-2},BOD_{-3}	BOD_1	10.4178	18.0067	100	0.2
MO3	$MO3_d$	6	$BO0_0$,BOD_{-1}, BOD_{-2},BOD_{-3}, BOD_4,BOD_{-5}	BOD_1	10.6227	18.2476	5000	0.2
MO4	$MO4_d$	8	BOD_0,BOD_{-1}, BOD_{-2},BOD_{-3}, BOD_{-4},BOD_{-5} BOD_{-6}, BOD_{-7}	BOD_1	10.9127	18.3354	5000	0.2
MO5	$MO5_d$	10	BOD_0,BOD_{-1}, BOD_{-2},$BOD_{-)}$, $BOD_{-4)}$,$BOD_{-5}$$BOD_{-6}$, BOD_{-7}, BOD_{-8}, BOD_{-9}	BOD_1	11.4392	18.9103	3000	0.2
MO6	$MO6_c$	12	BOD_0,BOD_{-1}, BOD_{-2},BOD_{-3}, $BOD_{-4)}$,$BOD_{-5}$$BOD_{-6}$, BOD_{-7}, BOD_{-8}, BOD_{-9}, BOD_{-10},BOD_{-11}	BOD_1	12.2287	19.0435	5000	0.2

KNUST

LR is learning rate.

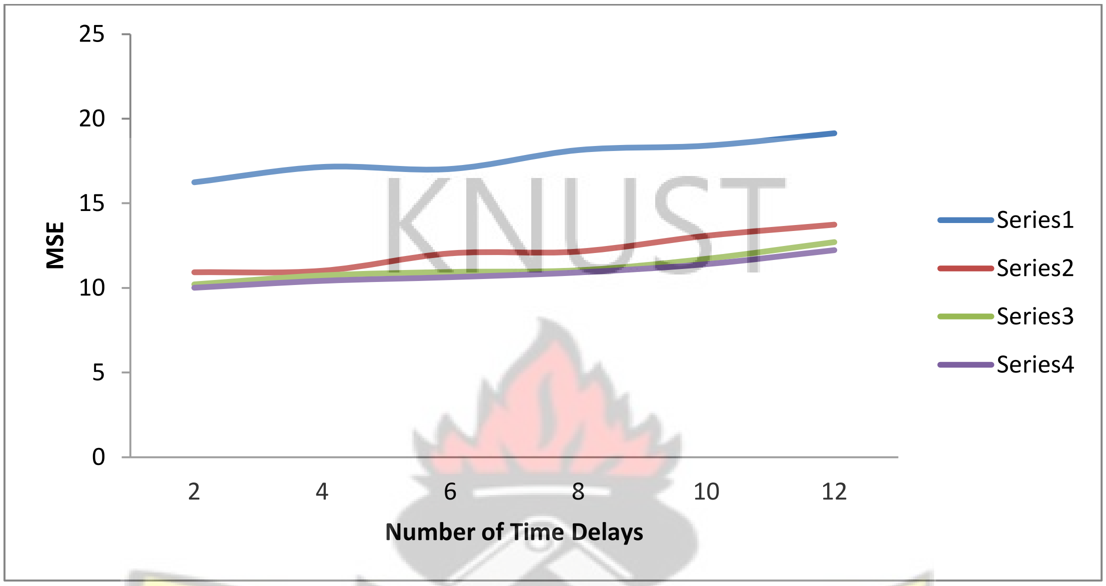


Figure 13 A Graph of MSE against Number of Time Delays

Serie1 represents the curve when the number of hidden neuron is 2 for all the six models

Series4 represents the curve when the number of hidden neuron is varied to 14 for all the six models

4.2 Training TDNN

The training of these models was started with the default values of neuralsolutions with a training count of 100 and 2 hidden neurons in the hidden layer. From the next trail, the optimum training count for the network was decided. This was done by trial and error by checking the MSE and the MAPE after each cycle of training. The optimum training count was the one which gave a minimum MSE and lower MAPE for the test data. After deciding the maximum training count for these models the numbers of hidden neurons in the hidden layer were varied by small increments by maintaining constant training count until the desired MSE and MAPE for the test data was obtained.

The training was done for these models by varying the learning rates of the network (0.2 to 0.75) and it was observed that there was no significant change in the MSE values after training. Therefore, all these models were trained with the default values of neuralsolution for learning rates. However, by varying the training count and the number of neurons in the hidden layer, the performance of the network greatly improved. The variation of MSE with different training count and hidden layers for model series are shown in Tables 8- 16 in the Appendix A and Appendix B. From these values it was observed that the MSE tends to cease after a particular time of training and almost remains constant throughout the training period.

CHAPTER FIVE

DISCUSSIONS

The variations of MSE with different training count and hidden neurons for the Models are shown in Appendix -B, Tables 9–14. After each set of training, MAPE for the test data was calculated. The various MAPE values obtained for the test data using these models are shown in Appendix- B, Table 15. The MSE performance indicates that lower error levels are achieved with a fairly large number of neurons in the hidden layer. There was a significant reduction in MSE from models with two hidden neurons to models with six hidden neurons as depicted in Figure 13 by the series 1 and series 2 respectively. There is also a meaningful MSE reduction as the number of hidden nodes increased from 6 to 10, and a further modest reduction going from 10 to 14 hidden nodes. These are also shown in Figure 13 as series3 and series4 respectively. It is also evident from Figure 13 that for any of the long term memory structures, there was a gradual increase in error as the number of short term memory delays increased. This is a fairly shallow linear relationship that is independent of the long term memory structure.

The optimal neural network time series model for forecasting BOD effluent is identified as a TDNN sub model; MO1d, with m = 2 and a hidden layer of 14 neurons. In the systematic analysis of TDNN performance, each of the six TDNN models was evaluated by MSE and MAPE and the results are displayed in Tables 9-14. It can be seen from the Table 9 that in MO1 the performance shows a decreased in MSE from 16.2441 of $MO1_a$ to 10.9237 of $MO1_b$. A further reduction in error is seen from 10.2071 of MO1c to 10.0088 of $MO1_d$. A similar reduction pattern was observed in all the sub models of the six models in this work. It was also observed that for each of the six models with 14 hidden neurons, there was a gradual increase in the error as the short term memory delay length increases. The MSE for $MO1_d$, 10.0088, increases to 10.4178 for $MO2_d$. As the short term memory delay length increases, the MSE error increases to 10.6227 for $MO3_d$, then to 10.9127 for $MO4_d$ and then a further increase to 11.4032 for $MO5_d$ with the final error reaching 12.2287for $MO6_d$. This trend was also seen across the models with the same number of hidden neurons. The results showed that using more time lagged series as the network inputs did not give much advantage. It was evident that the sub model, $MO1_d$, of Model MO1, with BOD_0, and BOD_{-1} as the input parameters was the best model for predicting BOD_1 with a MSE of 10.0088 in the training data and an MAPE of 11.6614. Model $MO1_d$ gave good results at a training count of 5000 and 14 hidden neurons in the hidden layer. All the other models showed comparatively poorer results than model $MO1_d$ with both training and testing. Time series neural network models are shown in this study to have the ability of forecasting wastewater effluent BOD, a complex application domain with nonlinear relationships and time varying dynamics.

The results of models obtained from NEURALSOLUTIONS collectively showed good statistical significance. Model $MO1_d$, was able to predict BOD_1 using BOD_0 and BOD_{-1} as the model inputs, at a training count of 5000 and 14 hidden neurons in the hidden layer. The results from this neural prediction showed very less MAPE values, indicating that the predictions are highly acceptable. Similar data driven modeling approaches can be developed to suit any industrial situation to predict fluctuating effluent concentrations well in advance.

CHAPTER SIX

CONCLUSIONS AND RECOMMENDATION

6.1 Conclusion

Wastewater treatment requires a more sophisticated forecasting approach because it is a complex process with highly variable inputs, nonlinear time varying dynamics, and an autocorrelated time series structure. While simple forecasting models can be used in forecasting applications, there is growing evidence that such methods lack the power to forecast wastewater BOD accurately. A more sophisticated model such as the TDNN predicts BOD levels better. This research used TDNN in the analysis of wastewater BOD. Results of the analysis reveal that predicting short term BOD levels based on a historical daily measurements is possible using a TDNN. The forecast accuracy of the neural network, as measured by MSE and MAPE, is significantly better. The research has therefore shown that the nonlinear TDNN neural network model has the accuracy to support the future development of advanced monitoring and control systems for wastewater treatment plants.

6.2 Recommendation

While the focus in this research was on one-step forecasts, multi-step forecasts may also be obtained from the neural network time series model. This can be accomplished by incorporating a feedback from the output of the network to the short term memory structure. This feedback would replace the external input, BOD_t during the generation of multi-step forecasts

Wastewater treatment plant managers may consider implementing a wastewater quality monitoring system with near real-time feedback control and a TDNN neural network forecast model. Quick turn-around BOD measurements can be fed to the neural network software system, providing a process analyst with a forecast of expected BOD values for one or more future periods. With the aid of a forecast monitoring strategy, the data analyst can decide whether or not to take action to adjust the process or to warn relevant officials of the potential for high effluent BOD. The result is a highly responsive system that facilitates more effective management of water quality and its impact on human health.

The results of this paper should be taken as preliminary and as a step towards combining neural networks and qualitative reasoning. Future research should include extensions regarding deriving knowledge about the behavior of systems evolving over time, considering other types of neural networks, improving the used algorithms, and others.

Pollution of water bodies in Africa and Ghana in particular is of serious concern to environmentalists. It is hereby recommended that such countries like Ghana should have waste water treatment plants at strategic points to help in the measurement and analyses of BOD levels in water.

REFERENCES

Abdel-Ghaffar (1988).*"Chemical Composition of Wastewaters*", in a study report in Alexandria and Giza, Egypt

Al-Salem (1987). "*Average Composition of Wastewater",* in a study report, Amman, Jordan.

Ayers and Westcott (1985). "*Parameter of Agricultural Significance*" in FAO Irrigation and Drainage Paper No 29 Rev. 1.

Bandyopadhyay, G. and Chattopadhyay, S. (2007)."*Single hidden layer artificial neural network models versus multiple linear regression model in forecasting the time series of total ozone",*Int. J. Environ.Sci. Tech.vol.4, no.1, pp.141−149

Berthouex, P. M. and Box, G. E. (1996)."*Time series models for forecasting wastewater treatment plant performance",*Water Resources, vol.30, no.8, pp.1865-1875.

Boger, Z. (1997). *"Experience in industrial plant model development using large-scale artificial neural networks"*, Information Sciences, no.101, pp. 203-216.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994)."*Time Series Analysis and Forecasting Control*." Englewood Cliffs, NJ: Prentice Hall Press.

Box, G. E. and Tiao, G. C. (1973)."*Intervention analysis with applications to economic and environmental problems",*Journal of the American Statistical Association, no.70, pp.70-79.

Chang, I., Tiao, G.C., and Chen, C. (1988).*"Joint estimation of model parameters in the presence of outliers"*.Technometrics, no.30, pp.193-204.

Cheek, S. and Wilkes, R. (1994)."*Monitoring processes using wireless data acquisition*", Water Engineering and Management, vol.141 no.10, pp. 17-18.

Chen, C. and Liu, L. M. (1993).*"Joint estimation of model parameters and outlier effects in time series*", Journal of the American Statistical Association, vol.88, no.421, pp. 284-297.

Coulibaly, P., Anctil, F., and Bobee, B. (2000).*"Daily reservoir inflow forecasting using artificial neural networks with stopped training approach*", Journal of Hydrology, vol, 230, pp.244- 257.

Darbellay, G. A. and Slama, M. (2000). "*Forecasting the short-term demand for electricity: Do neural networks stand a better chance*", International Journal of Forecasting, vol.16 no.1, pp.71-83.

De Vries, B. and Principe, J. C. (1992)."*The gamma model: A new neural network for temporal processing"*, Neural Networks, vol. 5, pp.565-576.

EPA, U.S.A., (1973)." *Monitoring Industrial Wastewater"*, Washington, D.C.

Feachem, S (1983). *"Possible Levels of Pathogens in Wastewater",* in a study report, Egypt.

Fletcher, R. and Reeves, C (1964)."*Function minimization by conjugate gradients*", Comput. J., pp. 149–154.

Gontarski, C. A., Rodrigues, P. R., Mori, M., and Prenem, L. F. (2000).*"Simulation of an industrial treatment plant using artificial neural networks*", Computers and Chemical Engineering, vol.24, pp.1719-1723.

Gunnerson,A. Shuval, A. and Arlosoroff, F (1984). *"Wastewater irrigation of vegetables and Ascaris prevalence in Darmstadt and Berlin, compared with other cities in Germany not practising wastewater irrigation*", in a study report, West Germany.

Hamed, M.M., Khalafallah, M.G., Hassanien, E.A. (2004). "*Prediction of wastewater treatment plant performance using artificial neural networks"*, Environ. Mod.Soft.vol. 19, pp. 919‒928.

Haykin, S (1998). "*Neural Networks: A Comprehensive Foundation",* Englewood Cliffs, NJ: Prentice-Hall.

Hobbs, B. F., Helman, U., Jitprapaikulsarn, S., Sreenivas K., and Maratukulam, D. (1998).*"Artificial neural networks for short-term energy forecasting: Accuracy and economic value",* neurocomputing, vol. 23, no.3, pp.71-84.

http:/www.fao.org/docrep/T0551E/t0551E/e04.htm#

KNUST

Hornell, J (1992).*"Industrial experiences with neural networks,"* in Proc. 6th Oklahoma Symp. Artificial Intell.,Tulsa, OK. pp. 127–133.

Hornik, K., Stinchcombe, M., White, H. (1989).*"Multilayer feed forward networks as universal approximators*", Neural Networks vol.2, pp.359−356.

Jorquera, H., Perez, R., Cipriano, A., Espejo, A., Victoria Letelier, M., and Acuna, G. (1998)."*Forecasting ozone daily maximum levels at Santiago, Chile*", Atmospheric Environment, vol.32 no.20, pp. 3415-3424.

Kandiah (199*0), "parameters used in the evaluation of agricultural water quality",* in a research report, pp 23-24

Kreesuradej, Wunsch, D. and Lane, M. (1994).*"Time-delay neural network for small time series data sets",* in Proc. World Congr. Neural Networks, San Diego, CA.

Lang, K. and Hinton, G. (1988).*"A time-delay neural-network architecture for speech recognition",* Carnegie-Mellon Univ., Pittsburgh, PA, Tech. Rep. CMU-CS-88-152.

Lee, T.L. (2004)."*Back-propagation neural network for long-term tidal predictions*", Ocean Eng. Vol. 21, 225−238.

Lindberg, C. F. (1997)."*Control and Estimation Strategies Applied to the Activated Sludge Process",* Ph.D. Thesis, Uppsala University, Finland.

Luk, K. C., Ball, J. E., and Sharma, A. (2001).“*An application of artificial neural networks for rainfall forecasting*”, Mathematical and Computer Modeling, vol.33, pp. 683-693.

Maier, H.R. and Dandy, G.C. (1998).*“The effects of internal parameters and geometry on the performance of back propagation neural networks: an empirical study”*, Environ. Mod. Soft. vol. 13, pp.193−209.

Maier, H.R., Morgan, N.,and Chow, C.W.K. (2004). *“Use of artificial neural networks for predicting optimal alum doses and treated water quality parameter”*, Environ. Mod.Soft.vol.19, pp. 485−494.

Makridakis, S. and Hibon, M. (1979). “*Accuracy of forecasting: An empirical investigation”*, Journal of the Royal Statistical Society, A, vol.142, pp. 97-145.

Metcalf and Eddy. (1995).“*Wastewater Engineering, Treatment, Disposal and Reuse”,* 5^{th} Edition, McGraw Hill, NY.

Morishita (1985). “*Effect of nitrogen-enriched polluted water can supply a considerable excess of nutrient nitrogen to growing rice plants*” in a study report

Parkinson, G. (1998).*”BOD measurement in real time*” Chemical Engineering, vol.105, no.11, pp. 25.

Poch, M., Bejar, J., and Cortes, U. (1993).*“Wastewater treatment data source information”,* Universitat Autonoma de Barcelona and Universitat Politecnica de Catalunya, Barcelona, Spain.

Poggio, T., and Girosi, F., (1990).*“Networks for approximation and learning”*, Proc.IEEE, vol.78, no.9, pp.1481−1497.

Prybutok, V. R., Yi, J., and Mitchell, D. (2000).*“Comparison of neural network models with ARIMA and regression models for prediction of Houston’s daily maximum* *ozone concentrations",*European Journal of Operational Research, vol.122,no.1,pp. 31-40.

Rummelhart, D.E., Hinton, G.E. and Williams, R.J. (1986)."*Learning representations by back-propagation errors*", Nature vol.323, pp.533−536.

Saad, E., Prokhorov, D. Wunsch, D. (1996).*"Advanced neural-network training methods for low false alarm stock trend prediction",* in *Proc. IEEE Int. Conf. Neural Networks,* Washington, D.C., PP. 2021-2026.

Sawyer C.N., McCarty P. L. and Parkin G.F. (1994)."*Chemistry for Environmental Engineering*", 4th Edition, McGraw-Hill International Editions.

Schuurmann, G. and Muller, G. (1994). *"Back-propagation neural networks recognition vs. prediction capability",*Environ. Toxicol. Chem. Vol.13, pp.743−747.

Shanno, D (1978), *"Conjugate gradient methods with inexact searches,"*Math. Operations Res.,vol. 3, no. 3, pp. 244–256.

Shuval, F. Pfudrisin D.and Gunsen, F. (1986),"*cattle grazing on fields freshly irrigated with raw wastewater, or drinking from raw wastewater canals or ponds*", in a study report from Melbourne, Australia and Denmark.,pp. 143-170

Spall, J. C. and Cristion, J. A. (1997), "*A neural network controller for systems with unmodeled dynamics with applications to wastewater treatment"*, IEEE Transactions on Systems, Man, and Cybernetics-Part B Cybernetics, vol.37, no.3, pp.369-375.

SPSS 16.0 SPSS Inc., Chicago, USA

Tsoukalas,T. and Uhrig, G.,(1997), "*Artificial Neural Networks and the Human brain*", a research report, pp 23-25

Tan, H., Prokhorov, D. and Wunsch, D. (1995).“*Probabilistic and time-delay neural-network techniques for conservative short-term stock trend prediction*”, in Proc. World Congr. Neural Networks, Washington, D.C.

UN Department of Technical Cooperation for Development (1985). “*Major constituents of typical domestic wastewater”,* in a research report.

Vauder, P (1994).*“Minimization methods for training feed forward neural networks*”, Neural Networks, vol. 7, no. 1.

Waibel, A (1989).“*Modular construction of time-delay neural networks for speech recognition*”, Neural Comput., vol. 1, no. 1.pp.36-46.

Waibel, A., Hanazawa, T., Hinton,G., Shikano, K. and Lang, K.(1989).*“Phoneme recognition using time-delay neural networks,*” IEEE Trans. Accoust., Speech, Signal Processing, vol. 37, no. 3, pp. 328–339.

Wan, E. (1993). *“Time series prediction by using a connectionist network with internal delay lines”,* in Time Series Prediction: Forecasting the Future and Understanding the Past,A. S.

Weigend and N. A. Gershenfeld, Eds. Reading, MA: Addison-Wesley, PP.197-217.

WHO (1984).“*Organic and inorganic constituents of drinking water of health significance”* in a research report.

Wen, C. H. and Vassiliadis, C. A. (1998).*“Applying hybrid artificial intelligence techniques in wastewater treatment”*,Engineering Applications of Artificial Intelligence, vol.11, pp. 685-705.

West, D., Dellana, S., and Jarrett, J. (2002).“*Transfer function modeling of processes with dynamic inputs”,*Journal of Quality Technology, vol.34, no.3, pp.312-323.

Weyerhaeser, M.(1990) "*Effectiveness of standard methods for the examination water and wastewater*", a study report, vol.23, no.5.pp.56-70

www.dtreg.com

Yeung, I. and Yung, Y. (1999)."*Intervention analysis on the effect on two sewage screening plants on the BOD_5 levels in the Hong Kong Victoria harbor and vicinity",*Environmental Monitoring and Assessment,vol.58, pp.79-93.

Zhang, G., Patuwo, B., and Hu, M. Y. (1998), *"Forecasting with artificial neural networks: The state of the art"*,International Journal of Forecasting, vol.14, no.1, pp. 35-62.

Zhu, J., Zurcher, J., Rao, M., andMeng, M. Q. H. (1998),"*An on-line wastewater quality prediction system based on a time-delay neural network*",Engineering Applications of Artificial intelligence, vol.11, pp.747-758.

KNUST
NYANSAPɔ WɔSANE NO BADWENIMA

APPENDIX A:

Time-Delay Neural Networks Used

In this section, the specifications of the considered network models are given. For all models, the inputs, the number of hidden neurons *#HN*, and the learning rate *LR* in the Time-Delay Neural Networks are given. The Biochemical Oxygen Demand is used as input.

An index i for $i \in \{\ldots, -2, -1, 0, 1,\}$ is used to indicate whether the value is given for today (1), yesterday (0), the day before yesterday (-1), another day before (−2) and so on. Variation of Number of hidden neurons from 2, 4, 10 to 14 gives corresponding sub-models for each model as MO_a, MO_b, MO_c, to MO_d respectively. All models predict the Biochemical Oxygen Demand (BOD_1) value for today.

Table 8 Models with their corresponding input parameters

m	Model	Input parameters	Output	#HN	LR
2	M01	BOD_0, BOD_{-1}	BOD_1	2,6,10, 14	0.2,0.7
4	M02	BOD_0,BOD_{-1},BOD_{-2},BOD_{-3}	BOD_1	2,6,10, 14	0.2,0.7
6	M03	BOD_0,BOD_{-1},BOD_{-2},BOD_{-3},BOD_{-4},BOD_{-5}	BOD_1	2,6,10, 14	0.2,0.7
8	M04	BOD_0,BOD_{-1},BOD_{-2},BOD_{-3},BOD_{-4}, BOD_{-5},BOD_{-6},BOD_{-7}	BOD_1	2,6,10, 14	0.2,0.7
10	MO5	BOD_0,BOD_{-1},BOD_{-2},BOD_{-3},BOD_{-4}, BOD_{-5},BOD_{-6},BOD_{-7},BOD_{-8},BOD_{-9}	BOD_1	2,6,10, 14	0.2,0.7
12	M06	BOD_0,BOD_{-1},BOD_{-2},BOD_{-3},BOD_{-4}, BOD_{-5},BOD_{-6},BOD_{-7},BOD_{-8}, BOD_{-9},BOD_{-10},$B0D_{-11}$	BOD_1	2,6,10, 14	0.2,0.7

APPENDIX B:

MSE and MAPE Errors

The following tables show the RMS error of all models of all test series. Training Count gives the necessary number of learning steps for the neural network in order to obtain the best result.

KNUST

Table 9 Variation of MSE with different training count and hidden neurons for Model MO1

Training Count	$MO1_a$	$Mo1_b$	$Mo1_c$	$Mo1_d$
100	16. 7275	11.2594	11.42583	11.0686
500	16.6842	11.1088	10.3699	10.1814
1000	**16.2441**	11.0786	**10.2071**	11.0480
3000	16.6841	**10.9237**	10.9133	10.4859
5000	16.6844	11.0939	11.1065	**10.0088**

Table 10 Variation of MSE with different training count and hidden neurons for Model MO2

Training Count	$MO2_a$	$MO2_b$	$MO2_c$	$MO2_d$
100	17.2282	11.4798	**10.7356**	**10.4178**
500	**17.1459**	**11.0129**	11.8875	11.9251
1000	17.2676	11.1399	12.0099	11.0722
3000	17.4343	11.4363	12.1321	11.2919
5000	17.2866	11.5890	12.1355	11.3855

Table 11 Variation of MSE with different training count and hidden neurons for Model MO3

Training Count	$MO3_a$	$MO3_b$	$MO3_c$	$MO3_d$
100	17.7696	13.0077	12.5474	12.7423
500	17.6355	12.5827	11.1635	11.2658
1000	**17.0274**	12.5422	11.0748	11.1323
3000	17 .2261	12.1407	11.1226	10.9521
5000	17.5293	**12.0408**	**10.9401**	**10.6327**

Table 12 Variation of MSE with different training count and hidden neurons for Model MO4

Training Count	$MO4_a$	$MO4_b$	$MO4_c$	$MO4_d$
100	18.6879	13.6312	11.6951	11.7402
500	19.2986	**12.1479**	**11.0402**	11.4483
1000	19.4269	13.6278	11 .9693	11.8553
3000	**18.1433**	13.8973	12.7534	11.4287
5000	19.3366	13.0859	11.9557	**10.9127**

Table 13 Variation of MSE with different training count and hidden neurons for Model MO5

Training Count	$MO5_a$	$MO5_b$	$MO5_c$	$MO5_d$
100	19.2018	**13.0168**	**11.7182**	11.5014
500	**18.4038**	13.3759	11.8844	11.6142
1000	19.4597	13.7479	11.9748	11.5846
3000	18 .9919	13.9430	12.0353	**11.4032**
5000	19 .0421	13.8852	12.3596	11.7984

Table 14 Variation of MSE with different training count and hidden neurons for Model MO6

Training Count	$MO6_a$	$MO6_b$	$MO6_c$	$MO6_d$
100	**19.1380**	14.6112	12.8169	12.2704
500	20.3983	**13.7379**	12.9402	12.3392
1000	20.0239	14.6211	**12.7057**	12 .8433
3000	19.9833	14.1773	13.3754	12.4654
5000	19.6366	14.0859	12.9121	**12.2287**

Table 15 MAPE values for the models

MO1	MAPE	MO2	MAPE	MO3	MAPE
$MO1_a$	22.9312	$MO2_a$	22.9213	$MO3_a$	22.7321
$MO1_b$	18.6231	$MO2_b$	18.5214	$MO3_b$	19.3321
$MO1_c$	17.8342	$MO2_c$	18.1234	$MO3_c$	18.3321
$MO1_d$	**17.7233**	$MO2_d$	**18.0213**	$MO3_d$	**18.2213**

Table 15 MAPE Values for the Models, Continued

MO4	MAPE	MO5	MAPE	MO6	MAPE
$MO4_a$	23.5432	$MO5_a$	23.8324	$MO6_a$	24.9321
$MO4_b$	19.1321	$MO5_b$	20.0213	$MO6_b$	20.6321
$MO4_c$	18.4432	$MO5_c$	19.2231	$MO6_c$	19.5432
$MO4_d$	**18.3123**	$Mo5_d$	**18.9023**	$MO6_d$	**19.0234**

KNUST

Table16 Plotted Values of Error Verses Number of Time Delay

	Error	Error	Error	Error	m
Model	a $(\# \text{HN} = 2)$	b $(\# HN = 6)$	c $(\# HN = 10)$	d $(\# HN = 14)$	
MO1	16.2441	10.9237	10.2071	10.0088	2
MO2	17.1459	11.0129	10.7356	10.4178	4
MO3	17.0274	12.0408	10.9401	10.6327	6
MO4	18.1433	12.1479	11.0402	10.9127	8
MO5	18.4038	13.0168	11.7182	11.4032	10
MO6	19.1380	13.7379	12.7057	12.2287	12